CPSL

Proceedings of the Conference on Production Systems and Logistics

CPSL 2022

17th May 2022 – 20th May 2022
University of British Columbia
Vancouver, Canada

by
publish-Ing.

Partners of the Conference

WGP Wissenschaftliche Gesellschaft für Produktionstechnik
TIB Leibniz-Informationszentrum Technik und Naturwissenschaften Universitätsbibliothek
UBC The University of British Columbia
WZL RWTH Aachen University
TUM iwb - Institut für Werkzeugmaschinen und Betriebswissenschaften
IFA Institut für Fabrikanlagen und Logistik
Universität Stuttgart Institut für Industrielle Fertigung und Fabrikbetrieb IFF
PTW TU Darmstadt
Institut für Werkzeugmaschinen und Fertigungstechnik IWF
fir an der RWTH Aachen
Fraunhofer IGCV
Institut für Technologiemanagement Universität St.Gallen
LPS Lehrstuhl für Produktionssysteme
match Institut für Montagetechnik Prof. Dr.-Ing. Annika Raatz
UBC Sauder School of Business

www.cpsl-conference.com

CPSL

Conference on Production Systems and Logistics

International Conference, CPSL 2022, hosted at the
University of British Columbia in Vancouver, Canada
17th May 2022 – 20th May 2022
Proceedings

Editors:
Dr.-Ing. David Herberger
M. Sc. Marco Hübner

DOI: https://doi.org/10.15488/12314
ISSN: 2701-6277

publish-Ing.

Herberger & Hübner

Senator-Burda-Straße 31a,
77654 Offenburg
Germany

E-mail: info@cpsl-conference.com
USt-IdNr.: DE324911115

Conference Website www.cpsl-conference.com

If contents of the proceedings are used for further work, these are to be clearly referenced in accordance with good scientific practice. The recommended citation is:

'Name 1, Given Name 1'; 'Name 2, Given Name 2';'…': 'Titel'. In: Herberger, D.; Hübner, M. (Eds.): Proceedings of the Conference on Production Systems and Logistics (CPSL 2022), 2022, publish-Ing., S. XX-XX.

Foreword

It is our great pleasure to publish the proceedings of the third Conference on Production Systems and Logistics (CPSL 2022), which was hosted at the University of British Columbia (UBC) in Vancouver, Canada.

Finally… after two long years of COVID-19 and mainly meeting online we were able to meet again in person – and it was great. With more than 80 participants, we enjoyed an unforgettable conference at the UBC, one of the most renowned universities in the world – and on certainly one of the most beautiful campuses in the world.
However, it was anything but clear that the conference could take place on site. The first two conferences were already strongly influenced by COVID. At the first conference in March 2020 in Stellenbosch, the pandemic was just getting underway. Nevertheless, not even half of our participants could be there on site. Also, at the second conference in August 2021, unfortunately, it was simply unthinkable to meet everyone in person. For this reason, the conference was hosted digitally in an innovative model.
We are now all the happier that we were able to hold the conference on site, as it was always originally planned, in order to better meet the purposes of the CPSL. We founded the conference to provide a platform for young scientists to network and exchange ideas. A platform where we can discuss exciting research results and exchange views on current topics with like-minded people. And, of course, a platform where scientists can easily upload their papers in the field of production systems and logistics and where these are reviewed in a serious and comprehensive review process in order to continuously improve the quality of the papers together. With the CPSL 2022, we are confident that we have been able to take a big step towards what we want to achieve with the conference. The interesting presentations, the exciting discussions inbetween the sessions and of course the networking during and after the conference make us proud to be part of the CPSL and let us look forward to the next conferences with great excitement.

We would like to express deepest appreciation and gratitude to all our partners, without whom this conference would not have been possible. A special thanks goes to all the reviewers for engaging in this demanding but also encouraging review process and to our keynote speakers for their inspiring and interesting presentations. Finally, we would like to thank all our participants and fellow researchers who have been willing to share their research knowledge and experience with all of us and made the CPSL 2022 such a great event.

We are looking forward to meeting you again at the CPSL 2023.

Dr.-Ing. David Herberger
Conference Chair

M. Sc. Marco Hübner
Conference Chair

Review Process

The Conference on Production Systems and Logistics (CPSL) is an international forum for the scientific exchange on current findings in the field of production engineering.

For the submission of a paper, an abstract had to be uploaded considering the following main topics:

- Automation
- Business Administration
- Digitalization & Industry 4.0
- Factory Planning
- Knowledge Management
- Lean Manufacturing
- Machine Learning & Data Mining
- Production Planning & Control
- Process Management
- Supply Chain Management
- Sustainability
- Technology Driver

The submitted abstracts were evaluated in an internal review process, whereby successful submissions were invited to upload a full paper. Full papers had to adhere to a specific template and format provided on the CPSL website.

The submitted full papers were reviewed in a two-stage peer review process by experienced scientists from renowned research institutions as well as authors of other submitted papers. This process ensures a constant and high quality as well as the influence of all participants on the papers and reviews. Consequently, each paper submitted was sent to at least two reviewers, with a third reviewer being requested in case of non-consensus between the first two reviewers.

The reviewers were asked to review the submitted papers on the basis of a provided evaluation template and were encouraged to give detailed comments and suggestions for improvement. Among others, the following key questions were considered:

- Does the title reflect the contents of the paper?
- How do you rate the comprehensibility and logical presentation of the paper?
- How do you assess the relevance and originality of the topic described?
- How do you assess the practical relevance of the paper?
- Do you consider the work a proof of a thorough research and knowledge of the latest literature in the field of research?
- Are the conclusions clear and valid?

After completion of the reviews, all authors were given sufficient time to improve their papers according to the remarks of the reviewers. For more information on the review process and the "Publication Ethics and Publication Malpractice Statement" please visit the conference website.

CPSL

Acknowledgements

Our sincere thanks goes to our outstanding supporters who made this great event possible.

Publishing

publish-Ing.

TIB
LEIBNIZ-INFORMATIONSZENTRUM
TECHNIK UND NATURWISSENSCHAFTEN
UNIVERSITÄTSBIBLIOTHEK

Scientific Societies

WGP
Wissenschaftliche
Gesellschaft für
Produktionstechnik

Reviews & Scientific Support

WZL | RWTH AACHEN UNIVERSITY

TUM
iwb - Institut für
Werkzeugmaschinen und
Betriebswissenschaften

IFA
Institut für
Fabrikanlagen und Logistik

Universität Stuttgart
Institut für Industrielle Fertigung
und Fabrikbetrieb IFF

PTW
TU DARMSTADT

fir an der
RWTH Aachen

Institut für Technologiemanagement
Universität St.Gallen

Fraunhofer
IGCV

Institut für Werkzeugmaschinen
und Fertigungstechnik IWF

UBC SAUDER
SCHOOL OF BUSINESS

LPS
LEHRSTUHL FÜR
PRODUKTIONSSYSTEME

match
Institut für
Montagetechnik
Prof. Dr.-Ing. Annika Raatz

Keynote Speakers 2022

A special THANK YOU goes to our outstanding Keynote speakers, who joined the conference onsite or online and inspired our participants beyond their presentations with cutting-edge and highly interesting topics

Prof. Dr.-Ing. Rüdiger Daub

Head of the chair of Production Technology and Energy Storage Systems, Technical University of Munich and head of Fraunhofer Institute for Foundry, Composites and Processing Technology IGCV

Emerging research questions in manufacturing technology arising from electromobility

The shift towards electric cars and the rapidly increasing demand for battery cells pose major challenges for manufacturing technology. The keynote will show how the use of methods like artificial intelligence and data analytics can be used to increase traceability, quality and productivity in battery production. It will also highlight how the introduction of new materials and solid-state batteries will impact production systems.

Prof. Dr.-Ing. Joachim Metternich

Head of the Institute of Production Management, Technology and Machine Tools, Technical University of Darmstadt

The high-performance value stream: genuine customer individuality in the shortest delivery time

The keynote will give an overview of high-performance value streams and how individuality is made possible without adaptation effort. The main topics will be the integration of the customer into the value stream and the enabling of value streams for the entire solution space.

Prof. Dr. Thomas Friedli

Professor of Production Management and Director at the Institute of Technology Management, University of St. Gallen

The Future of Production Systems

For decades the most admired examples when it came to manufacturing stemmed from the automotive industry. Companies like Toyota or later Porsche and Audi became role models for the implementation of lean. However, more recently lean deteriorated in dozens of companies to mere tool boxes neither really engrained in the culture nor in the structure and also the bottom line contribution to success can hardly be identified and described. The future will see a revival of once intuitively understood principles but now scientifically derived and fully integrated. The speech will highlight this development based on data and practical examples.

Table of Contents

Table of Contents

CPSL Table of Contents

Table of Contents

Table of Contents

Table of Contents

Table of Contents

3rd Conference on Production Systems and Logistics

Efficient Intralogistics Planning Based On An Innovative Intralogistics Tool Using The Example Of A Flexible Battery Cell Factory

Anja Munzke[1], Thomas Neuhäuser[1], Andrea Hohmann[1], Rüdiger Daub[1,2]
[1]*Fraunhofer Institut for Casting, Composite and Processing Technology IGCV*
[2]*Institute for Machine Tools and Industrial Management, Technical University of Munich*

Abstract

In the course of increasingly volatile markets, globalization as well as shorter product life cycles, factories and thus also the logistics system as a central component of a factory have to be designed in a more flexible way. Battery cell production faces a special challenge in this aspect. Due to the trend towards a sustainable and environmentally friendly energy supply and mobility, the demand is expected to increase significantly. New battery cell factories have to manage rising product volumes and simultaneously react versatile regarding new research findings. Thus, the market for battery cells, the product itself, and the corresponding manufacturing processes are constantly changing. New materials, manufacturing methods, variations of cell formats as well as the possibility of scalability and the associated changes in the requirements for the factory must be taken into account as early as possible in the planning stage. The logistics system as one of the core elements of a factory is always affected by changes in the product, manufacturing processes or input materials. If, for instance, other materials are used, the storage and transport of these goods with different dimensions, weight or even environmental requirements must still be guaranteed.

In order to consider the required flexibility already in the planning process, simulation can provide a decisive benefit. It enables the planner to analyse the production and iteratively adapt logistics planning. Since there are many possibilities and combinations, especially in the design of warehouse and transport systems, a reduction of these should take place at an early stage. However, the preselection of suitable logistics systems that provide the necessary flexibility is currently often based on empirical knowledge and extensive market research. Therefore, this paper presents an efficient, holistic approach to logistics planning and an intralogistics tool in detail, which is based on established data. As a result, an optimal logistics system can be defined through an iterative optimization of the flexibility corridor, taking into account the factory goals.

Keywords

Intralogistics; Flexibility; Battery Cell Factories; Simulation; Factory Planning

1. Introduction

In the course of the energy transition, all industries face new challenges in order to achieve the Paris climate targets. The automotive industry plays a major role to reach the targets. About 22% of the GhG emissions come from the mobility sector [1]. Electromobility is not just a trend anymore. The transformation from combustion engines to electric cars is forcing car manufacturers to act quickly. According to the IEA's Global EV Outlook 2021, there were around 10 million electric cars worldwide at the end of 2020, representing a

DOI: https://doi.org/10.15488/12148

growth rate of 41% in 2020 [2]. Under current conditions and assumptions, there will be approximately 145 million electric vehicles (excluding two/three-wheelers) on the roads worldwide by 2030, requiring a battery cell capacity of 1.6 TWh [2]. However, in order to be able to meet the climate targets despite changed framework conditions, around 230 million vehicles will have to be on the roads in 2030, resulting in a demand for battery cell capacity of 3.2 TWh. For comparison, the potential production capacity for battery cells worldwide in 2020 was 300 GWh [2].

Thus, the need for new battery cell factories will increase. Today, battery cells are already used in electric cars, but their development is not yet well advanced and variants that are more efficient are constantly being presented. Lots of research is being conducted into new, alternative materials, variations of cell formats as well as innovative manufacturing processes [3,4]. In addition, the demand for the available product variants is also changing. New battery cell factories must face these challenges.

One possibility is to design the factory highly flexible in terms of production volume, materials, product varaints and manufacturing processes. This flexibility should already be taken into account in the factory-planning phase [5]. In addition to the production system, the intralogistics system must be designed flexible. For example, the means of transport must be able to cover the transport capacity even if the production volume increases and must be suitable for transporting alternative materials [6,7].

Since the logistics system as part of the entire factory is very complex, planning requires methods and tools to manage this complexity. To support the intralogistics planning, simulations can be built up, which help to better understand the complexity of the logistics system and thus optimize and validate the planning [8]. The present work introduces a concept, how the required flexibility can be considered in simulation-based intralogistics planning, whereby simulation-relevant parameters are transferred over a developed tool into the simulation. The self-developed tool named Intratel is to be used to make a preliminary decision for possible storage and transport systems according to the required flexibility in battery cell production, so that the simulation is only build up with systems that provides the required flexibility and faster recommendations for action are possible.

Therefore, in chapter two the term flexibility in connection with factory planning is presented. Additionally, intralogistics or intralogistics planning while using a simulation in the context of factory planning is defined. Chapter 3 explains the concept with its individual components. Finally, in chapter 4, a short summary and conclusion is given.

2. Fundamentals

2.1 Flexibility as a target variable in the factory

Flexibility can be considered as a property of a system. It gives the system the ability to react to uncertainties and dynamics on the basis of specified action spaces, according to defined options for action. If a system is flexible, it can react and adapt to known but varying requirements [9–12]. In this context, a flexibility corridor defines the action space in which a system can adapt to changes in the environment for all possible future scenarios [10].

In the present work, flexibility shall exclusively refer to the meaning within a factory. One way to classify flexibility in the context of a factory is to divide it into three levels of consideration. According to Sethi and Sethi [13], component flexibilities, system flexibilities and aggregated flexibilities are differentiated. Subsequently, eleven types of flexibility are assigned to these levels. Relevant types for the present work are, product mix flexibility, which means the simplicity of introducing new products, program flexibility which describes stability of the system to produce different variants, quantity flexibility, which describes the ability to remain economical at different utilization levels and material flow flexibility, which describes the ability to produce different work pieces efficiently on any paths through the system [13]. The concrete effects that these types of flexibility can have in the context of battery cell production are presented in chapter 3.1.

Since much of the planning process depends on the defined flexibility or flexibility corridors, these should be integrated in an early phase of factory or logistics planning. Factory planning is based on different goals for the factory, which are derived from the company's strategy and general conditions. These goals are already worked out and defined at the beginning of a factory planning project. According to the guideline VDI 5200, flexibility is also mentioned here as a possible factory goal [14]. A company must be competitive and generate profit. For logistics, this results in performance, cost and quality targets. Examples of performance targets are the correct execution of orders or meeting the delivery times. Quality targets, on the other hand, are, for instance, the ability to deliver or the flexibility of the willingness to perform in the event of changes in requirements. For cost targets, the avoidance/reduction of inventories, the optimal use of infrastructure, the maximum utilization of load carriers or an efficient use of resources can be mentioned [15,16].

Flexibility as a factory goal is also important, but may conflict with the other goals. For example, high program flexibility adversely affects quality targets. In addition, ensuring a defined level of flexibility usually leads to higher production costs. It is important to include all target dimensions, flexibility as well as performance, costs and quality, in planning process. In order to be able to achieve an optimal planning result, the goals must also be prioritized depending on the application [17]. Furthermore, it is essential to define a suitable flexibility corridor for each type of flexibility, which is in line with the other factory targets according to the prioritization.

2.2 Intralogistics and Intralogistics planning

The intralogistics system can be divided into individual subsystems such as machine, warehouse, picking, supply, buffer, handling and transport system [15]. In the present work, the focus is on the planning of a flexible transport and warehouse system for battery cell production. If logistics planning is considered as a subarea of factory planning, it must be integrated into the factory planning process. Figure 1 shows the seven phases of factory planning according to the VDI guideline 5200.

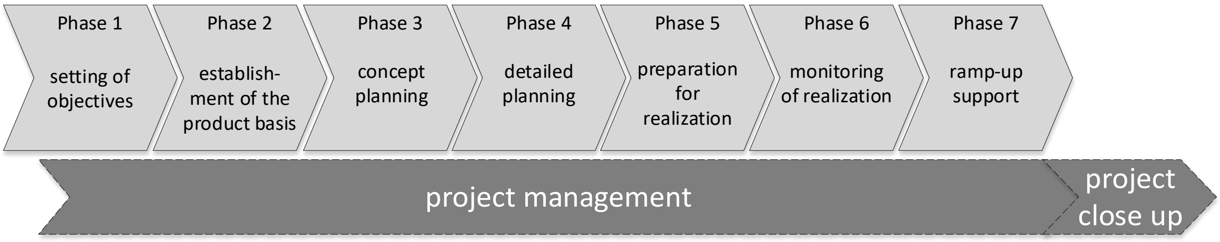


Figure 1: Factory planning phases according to the VDI guideline 5200

Logistics planning can be included in the conceptual and detailed planning phases, whereby prerequisites and boundary conditions for the subsequent planning phases are already defined in Phase 1 and Phase 2. According to the VDI guideline 5200, the concept phase (Phase 3) already results in a determination of the type and quantity of operating equipment and personnel resources, a dimensioning of logistics equipment, and the development of a material flow concept. In the subsequent detailed planning (Phase 4), the logistics concept is supplemented by process descriptions, whereby an assignment of products and resources to the respective processes becomes necessary [14].

The complex connection between factory planning (1) according to the VDI 5200 guideline and the logistics planning process (2) is detailed in Figure 2.

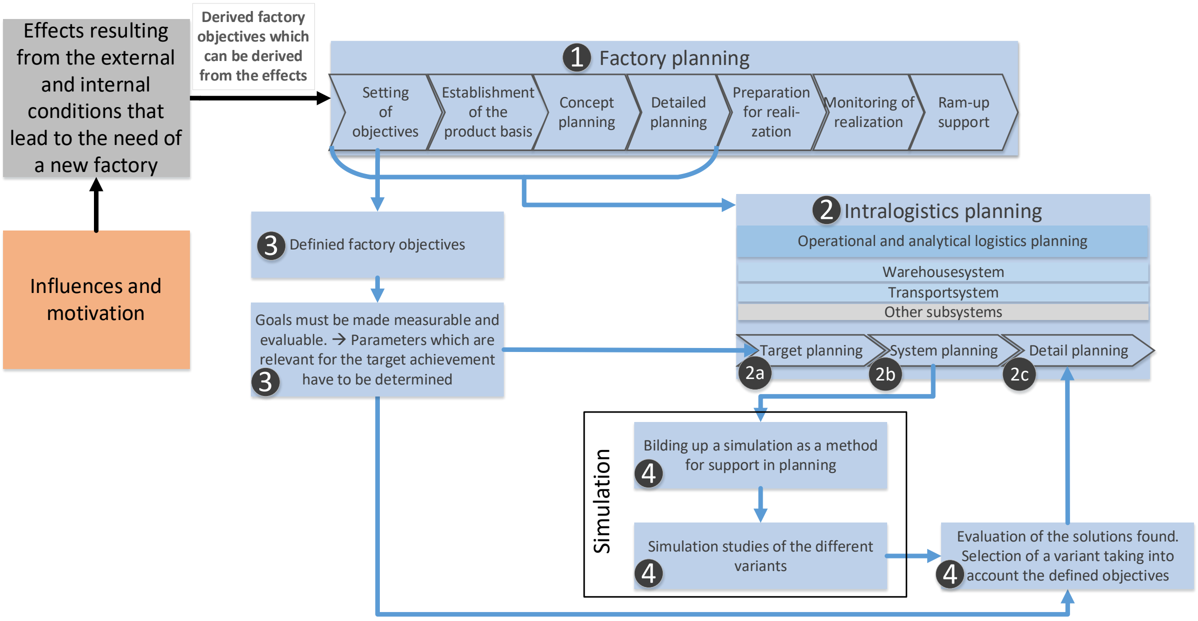


Figure 2: Factory and Intralogistics planning

The intralogistics planning can be divided into the phases target planning (2a), system planning (2b), and detailed planning (2c). Based on defined factory objectives (3), which were defined in the first phase of the factory planning process and need to made measurable and evaluable, all boundary conditions, the most important planning assumptions as well as performance requirements for the following phases are documented and, if necessary, analyzed in the target planning (2a). Thus target planning has a strong interaction with Phase 1 and Phase 2 of the overall factory planning process (1). In the next phase, the system planning phase (2b), possible solutions are developed, which also corresponds to the conceptual phase in the factory planning. In this context, suitable resources must be selected from a large number of possibilities. These are linked as performance points to logistics chains, organized and dimensioned so that specified performance targets can be met with specified restrictions as cost-optimally as possible. For the selection of a solution variant, methods such as a feasibility study or a performance comparison can be carried out. The feasibility study focuses on the technical feasibility, the fulfillment of the performance requirements, the compliance with the general conditions, restrictions and requirements and the feasibility within specified time. The performance comparison involves a crosscheck of marginal performance, capacity reserves, and flexibility of the selected solutions. In the case of flexibility, for example, it is considered whether the system is still suitable in the event of a change in the order structure or a change in logistics units or throughput and inventory fluctuations. In the profitability comparison, solutions are to be compared with each other with regard to their operating costs at maximum performance. Finally a utility value analysis is carried out, in which all remaining solutions are further compared with each other on the basis of defined criteria. In the subsequent detailed planning phase (2c), the selected solution is elaborated and getting prepared for approval. For this purpose, further specialist planners are consulted [15].

In system planning (2b) in particular, an optimum for the logistics system must be found on the basis of the different technical requirements and boundary conditions as well as the various factory targets. In principle, for this purpose many different methods and tools are available. However, analytical methods quickly reach their limits in a complex and dynamic system like a factory. In order to cope with the dynamic environment within the intralogistics planning, a simulation (4) can offer a considerable benefit. Often the interaction of boundary conditions and goals as well as dynamic behavior during future operation can be better understood if this is represented in a simulation model and different simulation scenarios can be executed. Multiple stochastic and dynamic interactions make it difficult to predict how a system will behave. Possible bottlenecks or potentials can be identified more quickly by means of a simulation and the logistics system can be planned and iteratively optimized accordingly [16]. The use of a simulation as a planning method and the current status of how intralogistics planning as part of the factory planning is done nowadays is shown in Figure 2. After the simulation is set up in the system planning phase (2b) and different variants are simulated (4), results must be evaluated. For this purpose, those parameters must be determined which are relevant for the achievement of the factory goals (3). After the evaluation, the selected variant enters the detailed planning (2c). However, the flexibility of the logistics system depicted in the simulation is already determined by the previous selection of technologies. An adjustment or optimization of the flexibility corridor, taking into account other factory goals, is therefore always associated with the manual search for suitable logistics technologies for the corresponding flexibility corridor before the simulation is set up. This complex process currently prevents the efficient mapping of different simulation scenarios with different flexibility coordinates and makes it difficult to achieve an optimum between all factory goals.

Simulation-based logistics planning is already known in the literature [18–20]. In the present work, a simulation is used as well. However, this simulation is coupled with a self-developed intralogistics TOOL (Intratel). The use of this tool makes it possible to select storage and transport systems early in the planning process that are suitable for the application and offer the necessary flexibility. Such a concept was not found in the literature so far to the best knowledge of the authors.

3. Efficient Intralogistics Planning For Flexible Battery Cell Factories Based On An Innovative Intralogistics Tool

This paper presents a concept to make the planning of a flexible logistics system more efficient. The construction of a battery cell factory is taken as a use case, since the requirement for flexibility is very important in future as described in section 1. Based on the conventional approach in logistics planning, the objectives for the logistics system in a flexible battery cell factory are derived (cf. Figure 3).

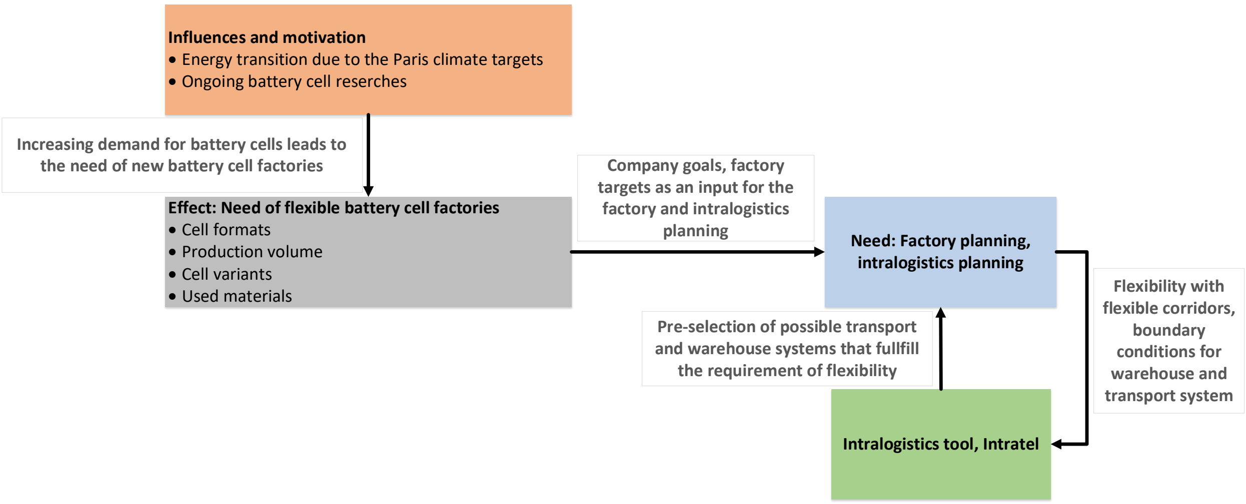


Figure 3: Concept for the intralogistics planning combined with the developed tool Intratel

In order to integrate the derived requirements in the intralogistics planning and to ensure an efficient planning process, the simulation model for the battery cell production is linked to a developed intralogistics tool, named Intratel. The goal is to get a pre-selection of possible storage and transport systems that meet the goal of flexibility by using the intralogistics tool in a first step.

This automatically leads to the fact that only systems are simulated, which can fulfill the goal of flexibility and can replace most of the components of a feasibility analysis as an evaluation method, while ensuring a holistic optimization of the logistic system. In the following, the overall concept with its individual components is presented. It includes the connection between the external influences (orange box), the flexibility requirements for a finished battery cell factory (grey box), the connected factory and intralogistics planning (blue box) and the developed intralogistics tool (green box).

3.1 Flexibility targets and relevant requirements

In addition to the typical factory goals, such as quality, performance and costs, battery cell factories must have a high degree of flexibility in the future (cf. chapter 1). The need for flexibility has various implications in the context of the logistics system of a battery cell production. In particular, the logistics systems must be able to cover the product mix flexibility, material flow flexibility, quantity flexibility and program flexibility.

Product mix flexibility

The introduction of new products should be considered in advance. In the case of a battery cell factory, this can be either a completely new cell format, such as round, pouch or prismatic cells or even solid state battery cells. Here, for example, the dimensions and weights of the finished products differ greatly. Also the input materials differ in some points. For round cells, for example, cans and lids are required, which have to be washed before they enter the clean and dry room of production. In the case of pouch cells, on the other hand, the case may first be deep-drawn during production and is supplied as a foil before. For each new product to be produced, it is necessary to estimate in which form and in which containers the input materials can arrive, so that suitable systems can be planned for both transport and storage. Another example for the extension of the product range could be electrode coils for which the storage and transport system would have to be designed, if they are part of the flexibility corridor.

Program flexibility

Alternative materials may be required for new variants, which may be supplied in other containers or on other loading equipment. An example of this is the solvent used in slurry production (first process step in battery cell production). Currently, NMP is often used for the cathode. This solvent is considered a hazardous material and is subject to special requirements for storage, handling and use. Alternative solvents are already being discussed. Storage and transport options should also be considered. If the factory also produces semi-finished products in the form of electrode coils, different variants can be produced here in terms of dimensions and weights. The storage and transport system used must be capable for these variants.

Quantity flexibility

If the factory is to be flexible in terms of quantity, the output quantity should be variable within the defined flexibility corridor. For example, the factory may be required to be able to produce between 1 GWh and 3 GWh per year. On the one hand, this requires more materials, which have to be stored, and on the other hand, the throughput has to be increased. Both have an impact on the transport system as well as the storage system. If material is stored or transported in larger quantities, for example, the use of tanks and silos for storage and transport via pipelines may be worthwhile. Alternatively, for smaller quantities, pallet storage can be used. The electrolyte and the active material can be mentioned here as example materials.

Material flow flexibility

In the context of battery cell production, the process steps slitting, calendering and vacuum drying can be interchanged within electrode production. Vacuum drying can also be an intermediate step in the assembly process. It may therefore be necessary to return the products to the vacuum dryers. Another requirement may be that, despite a separation into anode production line and cathode production line, both electrodes should be able to be produced on both lines. If required, the transport system used should be able to fulfill this flexibility within the material flow.

3.2 Intralogistics tool Intratel

The idea of the developed tool is to filter out from the multitude of warehouse and transport systems those that are suitable for the define flexibility corridor at hand before the simulation. Figure 4 shows the usage of the developed tool Intratel. To use the tool boundary conditions in particular to materials, semi-finished products, products and the production program must already exist.

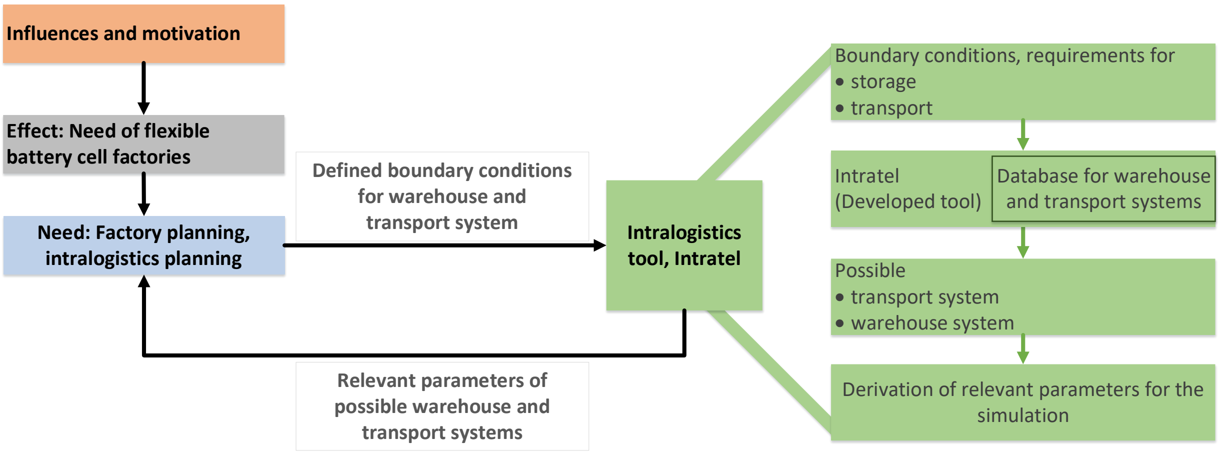


Figure 4: Intralocistics tool Intratel

The tool itself consists of a database, which has integrated all transport and warehouse systems with their properties and a user-friendly dashboard. Currently, the tool considers the selection of storage and transport systems separately. In the developed database, all fundamentally different transport and storage systems were summarized and classified. Also an initial division was made into continuous conveyors and discontinuous conveyors. In the next step, subcategories were formed within these categories, like belt conveyors or trackless floor conveyors. The final transport systems are then divided into these subcategories, such as apron conveyors, narrow aisle stackers or trackless AGVs. In the database for the warehouse systems are for example block storage, row storage, high rack storage for pallets, cantilever storage or carousels as options for warehouses. Furthermore, those characteristics were compiled, which can be relevant for the later selection. The properties were divided into two categories. These are especially important for the application. Mandatory attributes must be specified in order for the tool to make suggestions. An example of a mandatory attribute is the specification of whether bulk goods, general cargo or liquids are transported. Optional attributes can still be entered if they are already known in order to further reduce the number of possible systems. Currently, the tool works with 7 mandatory attributes and 16 optional attributes. The optional attributes include, for example, whether the quantity flexibility must be high. If there are too many requirements, some of them may be in competition with each other, which can lead to the fact that no suitable system is found.

The user himself sees only a dashboard on which he is guided through the attributes. For each attribute there are already predefined selection options. At the end, the user is shown those systems that are compatible with the requirements. Finally the parameters that affect the further factory targets, have to be integrated into the simulation model. In this application, the focus is on the corresponding parameters regarding capacities, personnel requirements, and the speeds of the individual systems.

3.3 Intralogistics planning based on a simulation and the tool Intratel

The developed concept shown in this paper uses a simulation as a method during the system planning phase for the intralogistics planning. The use of a simulation is not a mandatory method for planning. However, as shown in chapter 2.2, the dynamic and complex environment as well as the multitude of stochastic and dynamic interactions can be depicted more truthfully by a simulation and better statements can be made or the planning can be optimized. In the simulation model, the pre-selected warehouse and transport systems are now simulated and evaluated. If it was not possible to find systems that match on one hand with the required flexibility and on the other hand with all the other factory targets it could be that the flexible corridors have to be adapted and the pre-selection has to be done again. Afterwards one system for the warehouse and one for the transport system should be selected. This systems are then the input for the detailed planning phase of the intralogistics planning.

The simulation model was built with the software Tecnomatix Plant Simulation from Siemens. Modeling was done in 2D as well as in 3D. First, a network structure was developed for easier adaptability. Areas to be logically separated from each other were set up in individual networks. In cooperation with the manufacturing and production planning team, the networks for production were created. In addition to the production systems, all logistics areas such as provisioning, buffers, waste areas and workstations were integrated and properties such as capacity were assigned to these components on the basis of static calculations. An input table has been created so that all parameters can be changed and adapted for later simulation runs. This makes it easier to carry out planning and simulation scenarios. The planner can use this Excel table, which is linked to the simulation, to change the values before each run without having to go into the individual building blocks of the model. When starting the simulation, the current values of the table are applied. In addition, a warehouse concept including incoming and outgoing goods was developed in parallel, which was then mapped in a network. Delivery processes were integrated into methods and different delivery scenarios can be analyzed. No particular storage system has yet been selected. The first dimensioning regarding capacity was worked out on the basis of the production program and sales planning. Other important networks in the battery cell use case are locks, since these should not become a bottleneck in later production. Figure 5 shows a section of the simulation. In this model, production scenarios can already be run through and the effects on the utilization of the logistics elements used can be analyzed.

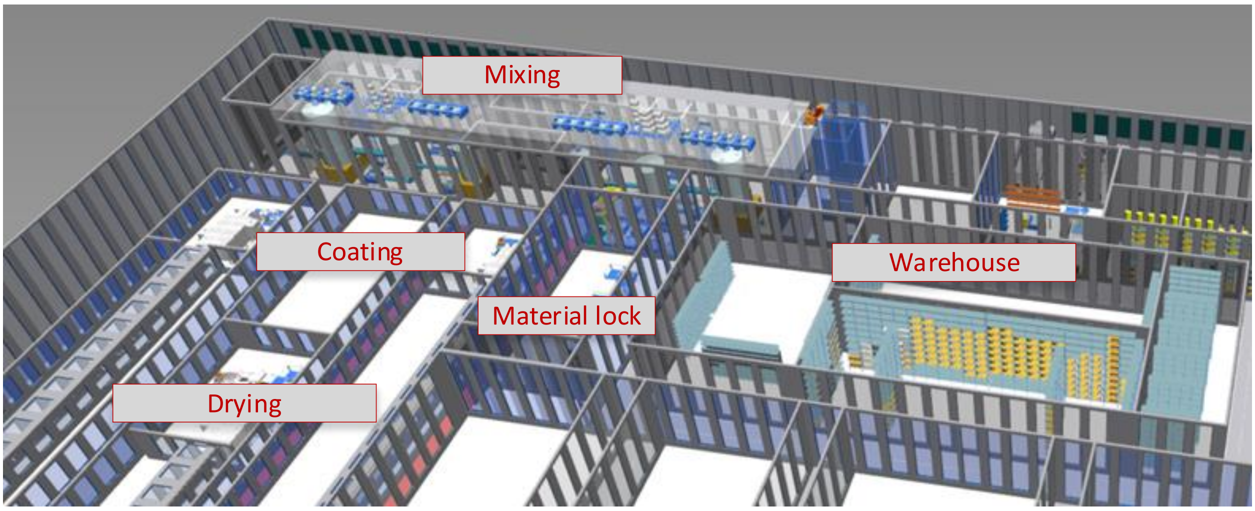


Figure 5: Section of the battery cell factory material flow simulation

In the present use case, flexibility was required in the context of product variants as well as a change of production volume. By using the tool Intratel, a pre-selection of possible transport and storage systems that can meet the flexibility requirements was made. As soon as the tool identifies suitable systems for the given application, those parameters are also determined which have an influence on the achievement of the factory goals. In the case of transport systems, for instance, the speed of the system. These parameters are in turn given as input to the simulation model.

4. Summary and conclusion

Intralogistics planning as part of factory planning requires structured and targeted implementation in every application. In the construction of a battery cell factory, planning is subject to many boundary conditions. New findings on production technologies, product variants and new materials are constantly generated and at the same time, factories must be scalable. The production and logistics system must be adaptable. Flexibility as a factory goal must therefore be integrated into the planning process. In order to support the planning process methodically and to be able to map the complexity, a simulation can be built up as in the developed concept. Within simulation studies it should be analyzed how far the factory goals can be reached with the plans. The developed tool Intratel can be used in advanced to get a pre-selection of possible warehouse and transport systems that can meet the flexibility requirements. Using the tool narrows down the warehouse and transport systems at an early stage of the planning. Here, the number of possible systems are suggested to the user on the basis of various input parameters as well as the flexible requirements.

In the next step, the concept shown must be verified in a concrete use case. The concrete application and subsequent evaluation should follow the developed concept. For this purpose, the interfaces between the simulation model, the tool and the corresponding evaluations must be worked out and implemented. In addition, the Intratel tool is to be further expanded to include additional subsystems for logistics planning.

5. Acknowledgements

Special thanks go to the German Federal Ministry of Education and Research (BMBF) for funding the research project "FoFeBatt" | Battery cell research factory. Furthermore, I would like to thank my colleagues Bernhard Voringer, Yash Visavadiya and Meltem Akkaya, who were significantly involved in the development of the simulation model.

References

[1] März A., 2019. Einfluss der Elektromobilität auf die Verteilnetze. edna conference, November, 2019.

[2] IEA, I.E.A. Global EV Outlook 2021.

[3] Hettesheimer, T., Thielmann A., Neef C., Möller K., Wolter M., Lorentz V., Gepp M., Wenger M. Prill T., Zausch J., Kitzler P., Montnacher J., Miller M., Hagen M., Fanz., Tübke., 2017. Fraunhofer Batterien Zellformate.

[4] Karlsruher Institut fuer Technologie, 2022.000Z. KIT - Das KIT - media - press information - archive press information - AgiloBat: Batteriezellen flexibel produzieren. https://www.kit.edu/kit/pi_2020_012_agilobat-batteriezellen-flexibel-produzieren.php. Accessed 22 March 2022.706Z.

[5] Abele E., Reinhart G., 2011. Zukunft der Produktion: Herausforderungen, Forschungsfelder, Chancen. munich: Hanser; 2011.

[6] Meier-Barthold, D., 2013. Flexibilität in der Material-Logistik. Springer; 2013.

[7] Schulte, G., 2010. Material-und Logistikmanagement, 2. ed. R. Oldenbourg publisher munich, wien; 2010.

[8] Bracht U., Schlegel M., Özkul F., 2019. "Fabrikplanung 5.0: Komplett virtuell zum Ziel" Zeitschrift für wirtschaftlichen Fabrikbetrieb vol. 114, pp. 208-212.

[9] Dormayer, H.-J., 1986. Konjunkturelle Früherkennung und Flexibilität im Produktionsbereich. munich, Univ., Diss., 1986. Ifo-Inst. für Wirtschaftsforschung, munich, 306 pp.

[10] Dürrschmidt, S., 2001. Planung und Betrieb wandlungsfähiger Logistiksysteme in der variantenreichen Serienproduktion. munich, Techn. Univ., Diss., 2001. Utz, munich, 186 pp.

[11] Pibernik, R., 2001. Flexibilitätsplanung in Wertschöpfungsnetzwerken. Deutscher Universitäts Verlag, Wiesbaden, 362 pp.

[12] Schmigalla, H., 1995. Fabrikplanung: Begriffe und Zusammenhänge, 1. edition ed. Hanser, munich, wien, 407 pp.

[13] Sethi, A.K., Sethi, S.P., 1990. Flexibility in manufacturing: A survey. Int J Flex Manuf Syst 2 (4), 289–328.

[14] Verein Deutscher Ingenieure, 02 2011. Fabrikplanung: Planungsvorgehen. Beuth.

[15] Gudehus, T., 2011. Logistik: Grundlagen - Strategien - Anwendungen, 4., updated edition 2010 ed. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg.

[16] März, L., 2011. Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Springer Berlin Heidelberg, Berlin, Heidelberg.

[17] Erlach, K., 2020. Wertstromdesign: Der Weg Zur Schlanken Fabrik, 3rd ed. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg, 430 pp.

[18] Golova, J., Mahmood, K., Raamets, T., 2021. Simulation based Performance Analysis of Production Intralogistics. IOP Conf. Ser.: Mater. Sci. Eng. 1140 (1), 12026.

[19] Mahmood, K., Karjust, K., Raamets, T., 2021. Production Intralogistics Automation Based on 3D Simulation Analysis. Journal of Machine Engineering, 101–115.

[20] Rohacz, A., Strassburger, S., 2021. The Acceptance of Augmented Reality as a Determining Factor in Intralogistics Planning. Procedia CIRP 104, 1209–1214.

Biography

Anja Munzke (*1993) is a research associate at the Fraunhofer Institute for Casting, Composite and Processing Technology IGCV and works since 2019 in the group of collaborative factory planning. She studied systems engineering with a specialization in logistics at the University of Applied Sciences THU in Ulm.

Thomas Neuhäuser (*1990) is group leader for collaborative factory planning at the Fraunhofer Institute for Casting, Composite and Processing Technology IGCV. He is chairman of the VDI-guideline committee 2552 Part 11.8 BIM – factory planning and spokesman of the buildingSMART Roundtable openBIM in factory planning.

Andrea Hohmann (*1983) studied aerospace engineering at the University of Stuttgart and received her doctorate in 2019 at the Technical University of Munich on the topic “Life cycle assessment of manufacturing processes for CFRP structures for the identification of optimization potentials”. Since 2012, Dr.-Ing. Andrea Hohmann is head of the department sustainable factory planning at the Fraunhofer IGCV.

Rüdiger Daub (*1979) is the holder of the Chair of Production Engineering and Energy Storage Systems at the Technical University of Munich and director of the Fraunhofer IGCV. After studying Electrical Engineering and Information Technology at the Karlsruhe Institute of Technology (KIT), he earned his doctorate at the Technical University of Munich (TUM). From 2012 to 2021, he held several positions at BMW Group Munich, most recently as Head of Technology Development and Prototype Construction of Lithium-Ion Cells.

3rd Conference on Production Systems and Logistics

Traceability System's Impact On Process Mining in Production

Markus Schreiber[1], Joachim Metternich[1]

[1] *Institute for Production Management, Technology and Machine Tools (PTW), Otto-Berndt Str. 2, 64287 Darmstadt, Germany*

Abstract

From the perspective of manufacturing companies, data handling is gaining more attention as it is becoming a strategic resource in digital ecosystems. Market forces such as rising amounts of product variants and decreasing batch sizes lead to higher complexity in manufacturing processes. Therefore, production management's demand for data-based process transparency is growing continuously as well as the number of companies turning to process mining to address these challenges. The increased use of process mining has uncovered many documented data quality issues that hamper output quality.

In response to data usage and quality problems, research in the field of Big Data has turned to sophisticated data value chains as a promising approach to optimize data usage. This paper presents the application of the data value chain concept on a manufacturing use case, delivering an assessment of traceability systems and their effect on data quality issues. This assessment reviews commonly known quality issues and investigates how traceability systems can influence and facilitate better data quality. The results support manufacturing companies in their use of traceability systems to improve the reliability of their process mining input data and, hence, their output performance indicators to meet the demand for more data-based process transparency.

Keywords

Data value chain; traceability system; process mining, data-based process transparency; data quality

1. Introduction

Industrial companies are challenged by rising complexity in their manufacturing processes. Good examples for complexity drivers are the constantly rising demand for more individualised products on the markets as well as short product lifecycles and delivery times. In response to competitors, companies create more product variants to stay attractive on their markets [2].

In the context of Industry 4.0, data-based transparency is needed to tackle complexity and support effective managers' decision-making [4]. To address process complexity, the use of process mining has become more popular amongst manufacturing companies. In a Deloitte study in 2020, out of 104 interviewed companies, 40% have stated to use process mining in production, aiming for process transparency and improvements as their top two reasons [6]. Still, the use of process mining is particularly challenging in production. Typically, production is characterized by numerous different processes that complicate identifying available data sources [8]. Available as well as reliable data are essential requirements for input datasets called "event logs" to conduct successful process mining analysis. Process mining projects tend to fail due various data quality issues such as missing and unreliable input data points that result in insufficient digital traces [9]. In his

DOI: https://doi.org/10.15488/12136

research, Jahn identifies data acquisition as the key factor for improving availability and reliability of input data and for gaining data-based transparency, the basis to optimizing production processes [10].

Within the vision of smart factories and smart products, automatic identification (autoID) technologies such as RFID are being used to generate data and gain transparency [13]. Although most manufacturing companies use these so called traceability systems due to legal obligations [14] or to inventory existing objects, the potential of data acquisition through autoID technologies is still not fully reached. Research in the field of traceability often focuses on tracking or tracing objects themselves, giving insights on how to consistently mark objects in production processes, find fitting technologies to track objects or products [14], or identify effort versus benefit levels that should be considered when tracking different product categories [16]. For successful production management, it has become crucial to focus on the data application perspective. A recent research project demonstrated several beneficial use cases originating from the use of a traceability system and its generated data [18]. In this context, companies still lack the knowledge to generate targeted feedback data of their processes using the traceability system and its ability to locate objects [4].

Traceability systems generate process data and can function as an important data supplier in production [16]. From a theoretical point of view, the combination of traceability as a data generating system and process mining as the tool for data analysis offers great potential for data-based process transparency [20]. However, researchers have not yet investigated whether and how traceability systems can avoid the occurrence of quality issues in input datasets. Based on a manufacturing use case, this paper aims to investigate the ability and impact of a traceability system to avoid common quality issues and improve the reliability of process mining outputs.

2. Approach

The paper is divided into two main sections. Section 3 addresses the conceptual development to identify data quality issues (QIs) that can potentially be affected by the traceability system in the manufacturing use case (section 3.3). This requires two tasks: Firstly, the explanation of the use case's data value creation process by introducing the data value chain (DVC) concept and the assignment of the use case to the phases of the DVC (section 3.1). Secondly, an overview about what kind of QIs exist and where the QIs occur in the DVC (section 3.2). Section 4 presents the analysis of the manufacturing use case. At first, the process mining input dataset (event log) and the obtained outputs based on the traceability data is introduced (section 4.1). Eventually, section 4.2 analyses the traceability system's impact to avoid the five most relevant QIs in the use case and to ensure reliable process mining outputs.

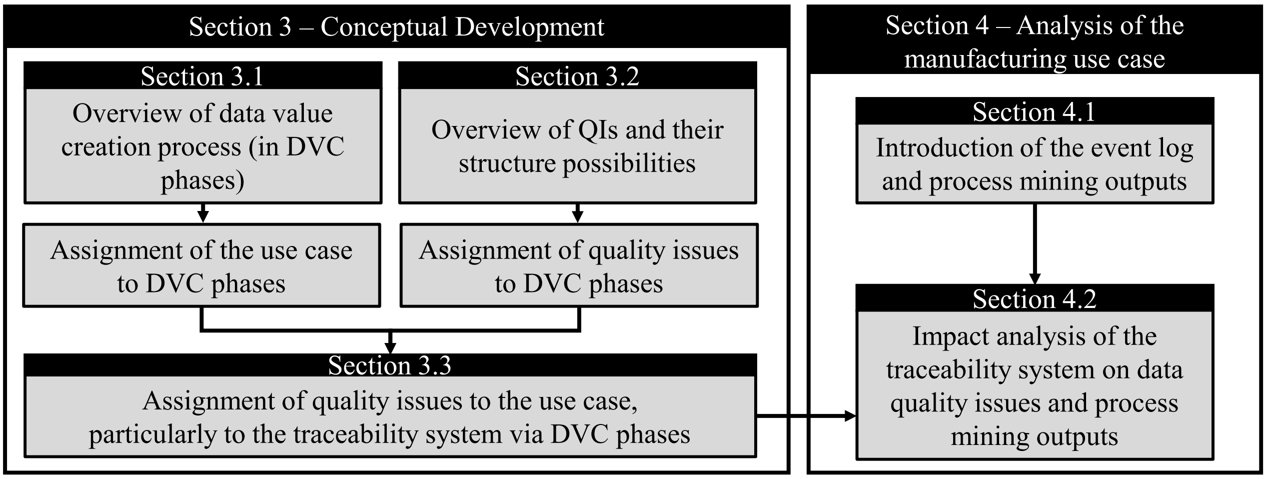


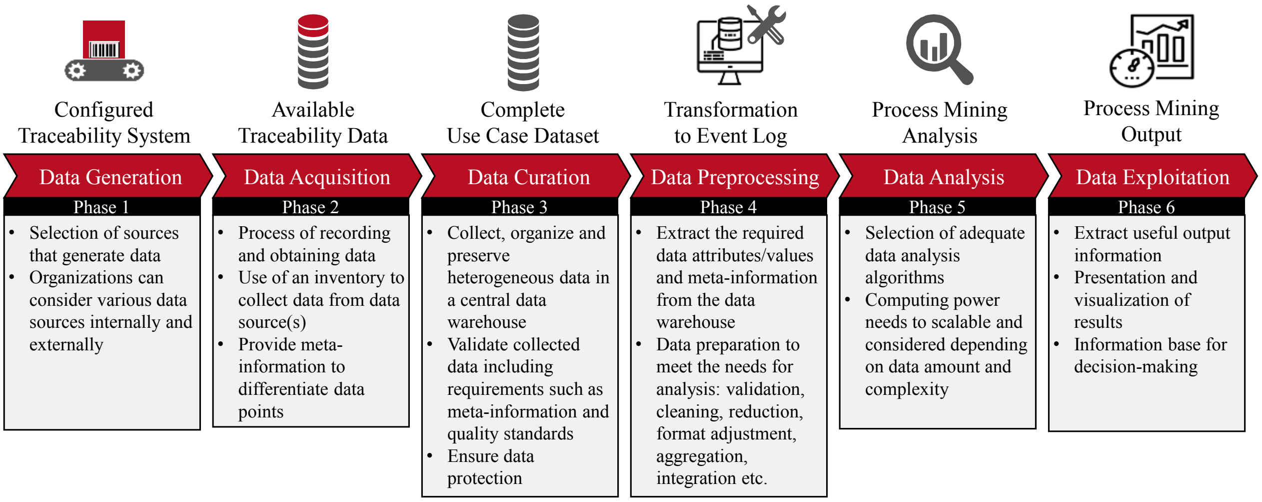
Figure 1: Approach of the conducted research

3. Conceptual development

3.1 Process of data value creation in the manufacturing use case

To enable the process of data value creation in companies the data value chain (DVC) concept is used. It represents a promising approach to improve the handling of data management. It originates from Big Data research and helps decision makers to adopt the data perspective on business processes in order to optimize data usage [12]. Generally, a DVC considers strategically important, value-creating activities [13] and integrates all data-effecting steps, starting with the generation and acquisition of data and ending with the possibility of decision-making based on data outputs [14,15]. Research shows that the representation of DVCs in literature differ in regard to the number of phases in the chain and the individually intended functions of each phase depending on the area of application [12,13,14,15,16].

To assess the impact of the traceability system on process mining in the manufacturing use case, the DVC is used to structure all relevant elements according to their role and function in the process of data value creation. These elements including the "configured traceability system", the "available traceability data", the "complete use case dataset" including traceability and more sensor data, the "transformation to event log" (filtered input dataset), "the process mining analysis" and finally "the process mining output". Figure 1 illustrates the application of the DVC concept on the manufacturing use case analysed in this paper. It suggests six phases that can be derived from the analysed sources considering the identified reoccurring patterns and functions explained in the grey boxes of each phase.



- Selection of sources that generate data
- Organizations can consider various data sources internally and externally

- Process of recording and obtaining data
- Use of an inventory to collect data from data source(s)
- Provide meta-information to differentiate data points

- Collect, organize and preserve heterogeneous data in a central data warehouse
- Validate collected data including requirements such as meta-information and quality standards
- Ensure data protection

- Extract the required data attributes/values and meta-information from the data warehouse
- Data preparation to meet the needs for analysis: validation, cleaning, reduction, format adjustment, aggregation, integration etc.

- Selection of adequate data analysis algorithms
- Computing power needs to scalable and considered depending on data amount and complexity

- Extract useful output information
- Presentation and visualization of results
- Information base for decision-making

Figure 2: Data value creation process based in the manufacturing use case

3.2 Literature review of typical data quality issues

The high quality of data is the key to interpretable and trustworthy data analytics and the basis for meaningful outputs. The process mining manifesto confirms this interdependency and stresses the need for high-quality event logs (representing the input data) in the context of process mining [3]. Typically, a company's decision to use an analytics tool such as process mining is made without paying attention to the possibly poor quality of the available event log which results in poor quality outputs – a dynamic often characterized by the term "garbage in – garbage out" [5].

The poor quality of event logs is a known problem amongst companies. A study by Suriadi et al. identifies several imperfection patterns from their experience with over 20 Australian industry datasets which confirm the severity of data QI in process data and their potential impact on process mining analysis [1]. Another study by the Meta Group revealed that 41% of the data exploration projects fail, mainly due to insufficient data quality, leading to misinformed decisions [7].

The goal of the conducted literature review is to identify typical QIs that can possibly be affected by the traceability system. The following criteria were applied in the literature search:

– A: Which literature source provide a collection of data QI?
– B: Do these QIs refer specifically to process mining?
– C: Do these QIs refer generally to data analysis?
– D: Are QIs structured in categories/dimensions based on commonalities?
– E: Are QIs structured by their location of origin?

Table 1: Literature Review

Source	[1]	[3]	[5]	[7]	[11]	[12]	[15]	[17]	[19]	[21]	[22]	[23]	[24]	[25]	[26]	[27]	[28]	[29]	[30]	[31]	Σ
A	•		•	•							•	•	•	•	•		•			•	**10**
B	•	•	•			•			•								•				**6**
C				•	•		•	•		•	•	•	•	•	•	•		•	•	•	**14**
D	•		•	•	•			•		•	•	•	•	•	•	•	•	•	•	•	**16**
E					•		•		•						•						**4**

Criteria A of Table 1 shows that there are several literature sources that provide collections of QIs. In literature sources related to process mining (criteria B) and general data analysis (criteria C), the amount of identified QIs is so large that a useful structure is required in order to tackle them systematically. Most studies (criteria D) provide a structure of QIs based on categories or dimensions. The general benefit of this structure is the fact that related QIs are gathered into the same group. However, this structure implicates a major disadvantage, as it hampers the search for the origin of the QIs. In contrast, the literature sources based on criteria E organize data QIs along the phases of the DVC as introduced in Figure 2. This approach has two main advantages, firstly it allows the identification of the QIs root cause by locating their places of origin [26], secondly, the structured QIs along the phases of the DVC can be used to connect the use case and its traceability system as it is linked to the DVC as well. It represents the basis to assess the traceability system's capability to avoid the occurrence of QIs.

Due to the large number of identified QIs, Singh et al. provide a study that summarizes a broad collection of data QIs acquired from extensive research in that field [26]. Based on the collected QIs, they suggest four groups of root causes for QIs that can be assigned to their places of origin in the phases of the DVC. The identified groups are the following [26]:

– **Group 1: QI based on data sources** – A leading cause for data QI is to obtain the wrong or poor data. On the one hand, every individual data source needs to be configured thoroughly to provide the data needed. On the other hand, various different data sources are likely to be inconsistent and cause difficulties in subsequent phases of the DVC.
– **Group 2: QI based on data profiling** – Once data sources are selected, the data profiling of every source system (e.g. traceability system, ERP, CRM, Web, etc.) needs to be examined to avoid negative impacts on data quality. The profiling is a fundamental step in which every individual source system as well as the gathered data of all source systems in a central data warehouse ensure data integrity and consistency for later analysis.
– **Group 3: QI based on data staging and ETL** (extraction, transformation, loading) – In this phase QI occur firstly in the central data warehouse when the data and metadata from all source systems is audited and validated and, secondly, in the pre-processing phase when a dataset is extracted, transformed and loaded for the following data analysis.
– **Group 4: QI based on data modeling** – If no major QI is detected up to this point and the available dataset demonstrates high quality, the data modeling itself can cause QI for two main reasons. The first

occurs when the dataset is not successfully transformed to fulfil the input specifications for the intended data analysis. The second can occur when the selection of the data analytics application is inadequate and obtains no or useless results.

3.3 Assignment of quality issues to the traceability system

The following step intends to determine which of the four identified groups of QI can be affected by the traceability system. Therefore, the groups of QIs mentioned above need to be linked to the DVC in the use case. Through the assignment of the identified groups of QIs and the manufacturing use case (see section 3.1) to the phases of the DVC, it is possible to link them as illustrated in Figure 2. This approach allows to break down the large amount of QI and assess a potentially positive impact of traceability systems on data quality and thus improve the process mining output.

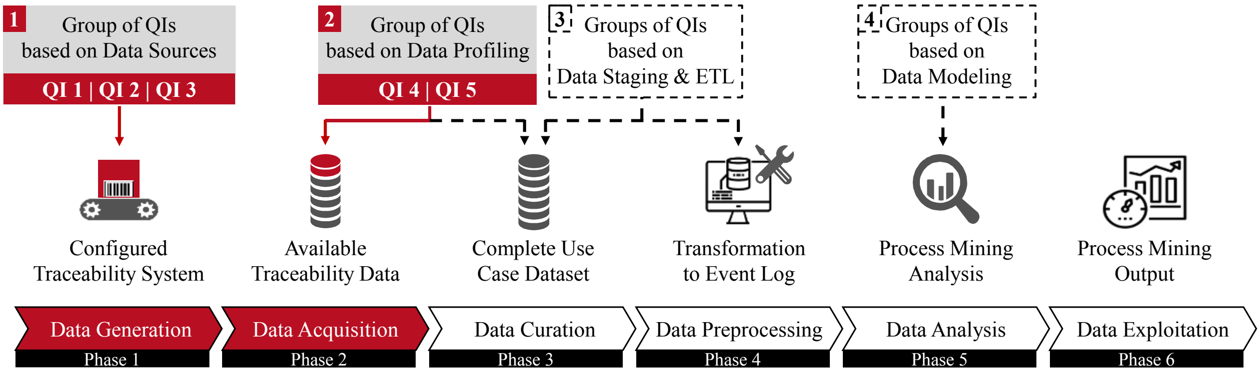


Figure 3: Link of QI groups to traceability system via DVC

Figure 3 demonstrates the result of connecting the groups of QIs as well as the elements of data value creation process in the use case to the DVC phases. As the research in this paper investigates the traceability system's impact on QIs, group 3 "QI based on data staging & ETL" and group 4 "QI based on data modelling" are not further considered in the analysis (shown black and white in Figure 3). There is a large number of identified single QI based on data sources (group 1) and data profiling (group 2) [26]. For handling purposes, the five most relevant QI (QI 1 – QI 5) for this use case are selected to explain the positive impact of the traceability system in connection to the conducted process mining analysis in section 4.2.

4. Analysis of the manufacturing use case

4.1 Introduction of the event log and the process mining outputs

The dataset in the use case was generated in the transfer project called "ArePron" (www.arepron.com). It represents a discrete production network involving parallel machine resources and consists of traceability data as well as sensor-based machine data (pressure consumption, electrical power consumption etc.). To investigate the traceability system's ability to prevent the emergence of QIs and to contribute to reliable process mining results, the machine sensor data is filtered and the traceability data providing process information remains to be used as input data for the process mining analysis.

The application of process mining requires a dataset as input data that contains at least a "case ID" (process trace) including "events" (process activities) and a "time stamp" for each event. Every case must be provided in a separate line [1]. The extracted traceability data from the dataset in this use case is shown in Table 2. Every individual "component No." functions as case ID, while the "machine name" and "process" as event, and "start scan" and "end scan" as time stamp. The event log is created in pre-processing (phase 4 of DVC), which mainly consists of format adjustments of the original dataset, so that any event (machining process) is given in a separate line.

Table 2: Use case event log with traceability data

Component No.	Machine name	Process	Start Scan	End Scan	[…]*
1042	Kasto	Sawing	08.03.2020 09:30	08.03.2020 09:38	
1042	OP10	Turning	08.03.2020 10:15	08.03.2020 10:42	
1042	OP20	Milling	08.03.2020 11:03	08.03.2020 11:23	
1042	[…]**				
1043	Kasto	Sawing	08.03.2020 09:40	08.03.2020 09:52	
1043	HaasST10	Turning	08.03.2020 10:31	08.03.2020 10:58	
1043	HaasMM2	Milling	08.03.2020 11:23	08.03.2020 11:48	
1043	[…]**				
1044	Kasto	Sawing	08.03.2020 11:00	08.03.2020 11:14	
1044	OP10	Turning	08.03.2020 11:27	08.03.2020 11:56	
1044	OP20	Milling	08.03.2020 12:29	08.03.2020 12:50	
1044	[…]**				

*Filtered data points (pressure consumption, electrial power consumption etc.) irrelevant for process analysis

**Following manufacturing steps in the same structure

The following represents a selection of exemplary process mining outputs shown in Figure 4 that are obtained from the traceability data:

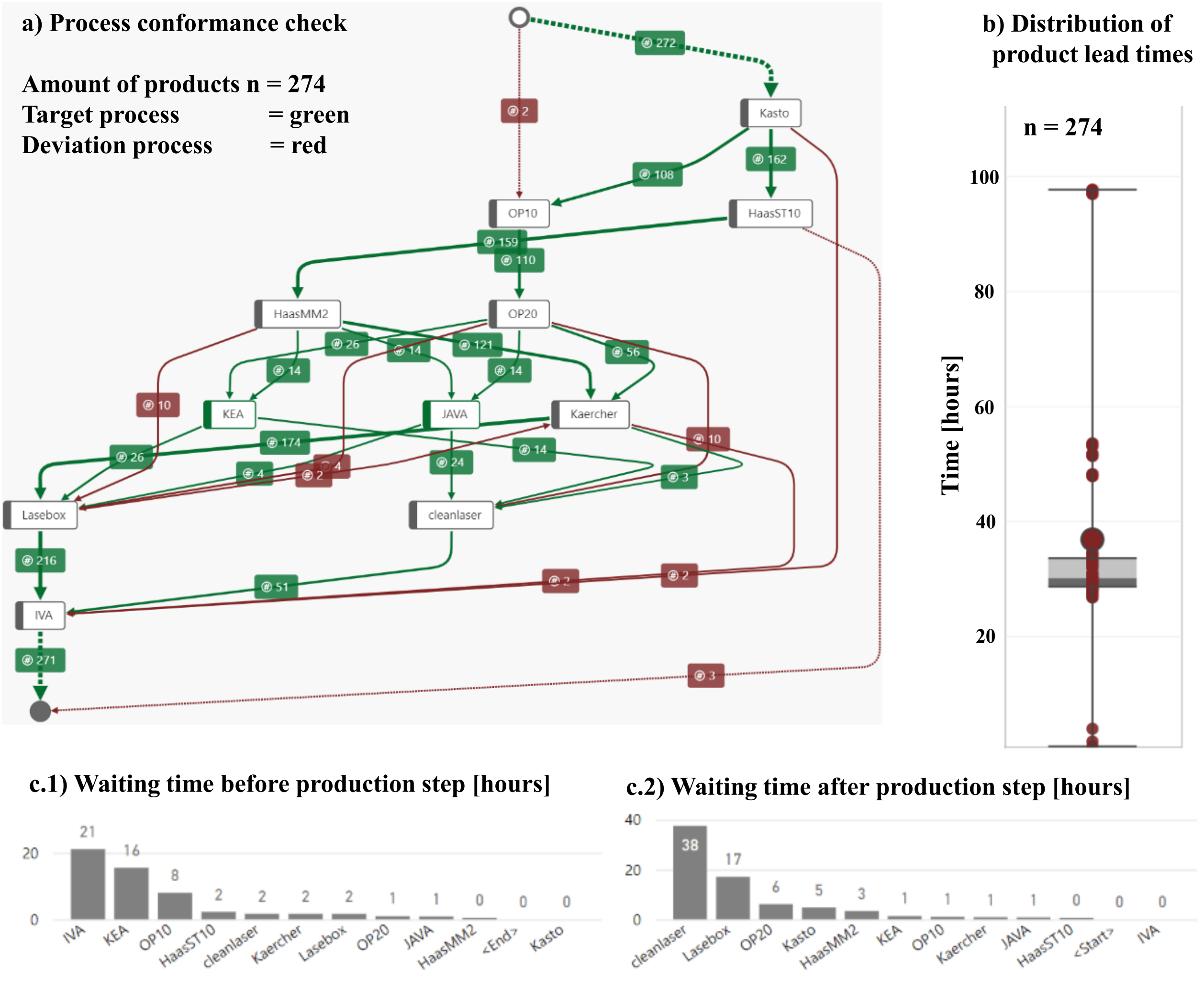


Figure 4: Exemplary process mining outputs

a) The standard process mining outputs process discovery and process conformance check are obtained. The conformance check in Figure 4 shows the amount of products that have taken the target process in green and those that deviated from their target process in red.
b) The available time stamp (either start or end time stamp) enable various output options regarding process lead time analyses, such as average lead time of all products, average lead time per production path, etc. The exemplary output distribution analysis of individual product lead times is shown in a boxplot in Figure 4. Each red point represents one product; the bigger the point, the more products were product in this lead time.
c) The availability of two time stamps (start and end time of each production step) allows for the analyses of the products' waiting times before and after every production step. Those indicators can hint at problems within production and can, for example, identify bottlenecks in the production network.

4.2 Impact analysis of the traceability system

To assess the traceability system's capability to avoid QIs and to ensure reliable process mining outputs, the impact analysis is conducted individually for every QI. At first, the analysis names the individual QI and outlines its relevance, then it explains the traceability system's impact on the QI and eventually on the process mining output (shown in Figure 4):

- **QI 1:** Inadequate selection of data sources do not provide needed input data

  Relevance of QI 1: The goal in the project is to analyse the performance of the production network and to receive transparency about the products' path within the production network, their lead time through the production process, wasteful waiting times etc. To receive the right process mining outputs, the selection of data sources is the most crucial success factor contributing to the project goal by improving the availability and reliability of input data and thus by gaining data-based transparency [10]. In case the planning and selection of needed source systems and their data points is neglected or has not been conducted at all, this QI is likely to result in missing analysis outputs. The effort to handle this QI is high, as the implemented hardware (source system, sensors etc.) needs to be changed and re-implemented.

  Impact of traceability system on QI 1: Generally, a traceability system's configuration determines the process data being captured, and what process mining outputs can be obtained. In this use case, the traceability system functions as the only source system to generate the process data shown in Table 2 and is responsible to ensure the performance evaluation by enabling the creation of the intended outputs. The traceability data required and generated are the "component ID" to capture the product, the "machine ID" to identify the taken paths through the production network, one "time stamp before the start" and another "time stamp after the end of a production process" (see Table 2).

  Impact of traceability system on process mining: The "component ID" functions as the case ID of the event log and represents the products trace through the production network. The "machine ID" functions as the event and determines the actual production stations passed by the product. The "time stamps" of component ID and machine ID captured in the production network help to order the different steps through production. As case ID, event and time stamp are the minimum requirement to create an event log, the exemplary process mining outputs such as process discovery, conformance check, lead times analysis or waiting time analysis shown in Figure 4 would not have been obtained without the traceability system.
- **QI 2:** Missing data values in data source system

  Relevance of QI 2: Depending on the implemented source system in each production case, the likeliness of missing some relevant data points in operational practice is high and therefore an important QI to be considered.

  Impact of traceability system on QI 2: Incomplete data generation with missing values is more likely to appear in a manually working traceability system than in a highly automated one. The ability to avoid missing data depends highly on the automation level of the traceability system and the operator's reliability in the case of a high manual level. In this use case a manual traceability system with hand scanners is implemented. An important measure for the successful and complete data acquisition is to instruct all users of the traceability system on its correct handling. Relying on the manual handling of the traceability system and a planned acquisition of 3.288 data points in 274 cases (sum of manufactured products), only 7 data points were missing. That's an error ratio of about 0.21%.

Impact of traceability system on process mining: Considering the error ratio of 0.21% in the event log, the process mining analysis and output is practically not affected by the few missing values.

- **QI 3:** Misspelled data in data source system

  Relevance of QI 3: Misspelled data points can occur especially when implemented source systems require manual system inputs by operators. Large amounts of misspelled data in a dataset may cause major efforts in the pre-processing phase when detected, otherwise subsequent data analysis become obsolete.

  Impact of traceability system on QI 3: Traceability systems offer technological possibilities such as optical labels (e.g. data matrix code) or RFID tags that save identification numbers and transfer those when captured via optical scanner or RFID receiver to a source system. In this use case, a data matrix code is used to provide the component ID and the machine ID so that misspelled data cannot occur during data acquisition.

  Impact of traceability system on process mining: Using the technological options to save relevant data in optical codes or RFID tags, no manual inputs into the traceability software are required. As result, there is no misspelled data available in the event log.

- **QI 4:** Insufficient data profiling of data sources such as lack of data validation routines at source system

  Relevance of QI 4: As Figure 3 demonstrates, the first opportunity to perform data profiling is possible at phase 2 (data acquisition) at the source system, such as the traceability system in this use case. The second opportunity is at phase 3 where a combined dataset is formed out of several possible source systems in a data warehouse. Validation routines represent data capturing rules that support the acquisition of the right data as needed. When applied in phase 2 and 3, the risk of crating QI is counteracted.

  Impact of traceability system on QI 4: In the use case, the traceability system is used for data profiling to ensure the high quality of the generated data. Therefore, data validation routines are embedded in the traceability system. This is even more important when the traceability system is operated manually and errors in the data generation phase are more likely to happen. For instance, the system captures only data points if the specified scan sequence is followed. For the traceability system to capture every individual production step of a product as valid data point, the machine ID must be scanned firstly, the component ID secondly. Additionally, both scans need to be performed within 10 seconds. This way the accidental capture of data can be avoided.

  Impact of traceability system on process mining: The used validation routines add significantly to the generation of a high quality event log. Component IDs or machine IDs are not obtained as individual data points so that the case ID and the corresponding event are always saved together. This way, outputs such as discovery, conformance check, lead times etc. are not distorted.

- **QI 5:** Hand coded data profiling is likely to be incomplete or results in an inappropriate selection of automated profiling rules

  Relevance of QI 5: QI 4 and QI 5 are related. QI 4 represents the conceptual level of what validation rules are useful to preserve the same data quality. QI 5 refers to the technical implementation of the validation routines via coding. Potentially, the selected validation routines embedded as rules in the system can declare correct data points as invalid and do not capture them.

  Impact of traceability system on QI 5: To ensure the coded rules in the traceability system contribute positively towards data quality and data completeness, tests with possible errors have been conducted to analyse if the implemented rules in the system function as expected and do not cause new QI. Moreover, the traceability system is designed to support operators by giving feedback if the intended scanning process is performed correctly and valid data is generated successfully. This gives system users the chance to verify if the system captures the correct data.

Impact of traceability system on process mining: The result of embedding coded rules for the automated validation of generated data in the traceability system is a reliable event log providing useful outputs as shown in Figure 4.

The analysis of the manufacturing use case demonstrates the high impact of traceability systems enabling the process mining analysis by generating the required process data. It outlines that the traceability system has the capability to improve the data handling by avoiding or at least minimizing the risk of QIs to occur and hence ensuring the reliability of the obtained outputs.

Summary and Outlook

This paper investigates the impact of traceability systems on data quality issues (QIs) and process mining results, based on a manufacturing use case in a production network. First, the connection between traceability and process mining is explained through the data value chain (DVC) concept in six phases. A thorough literature review results in the identification of four groups of QIs that are distinguishable by the location in which they occur along the DVC. Considering the application of the DVC phases on the use case, there are two (out of four) groups of QIs, "QIs based on data sources" and "QIs based on data profiling" that can be assigned to the traceability system and hence be positively affected by it. The investigation of five concrete QIs out of the two groups confirms that traceability systems can avoid QIs and improve the number and reliability of process mining outputs.

Traceability systems have a great potential to provide process data that enable transparency through process mining analysis in production. Due to growing complexity and more frequent use of process mining in production, traceability systems are not only relevant for commonly known purposes such as recall campaigns, but also take on an important role in today's data-based production management. As a supplier of valuable process data, they have the capability to enable transparency through process mining in production, firstly by providing the selected data points needed and secondly by its ability to prevent the occurrence of QIs.

Future research in the field of traceability needs to develop a scientific approach that allow companies the target use of traceability systems as a data supplier in production. This approach needs to address the question of how to configure a traceability system in order to maximises the number of process mining outputs and hence, the gained data-based transparency. At the same time, the traceability system's ability to avoid the potential occurrence of QIs needs to be considered in this approach, so that it contributes to reliable and high quality results of process mining analyses.

References

[1] Suriadi, S., Andrews, R., Hofstede, A. ter, Wynn, M.T., 2017. Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. Information Systems 64.

[2] Gottmann, J., 2019. Produktionscontrolling. Springer Fachmedien Wiesbaden, Wiesbaden.

[3] Van der Aalst et al., 2012. Process Mining Manifesto. International conference on business process management.

[4] G. Schuh, R. Anderl, R. Dumitrescu:A. Krüger:M. ten Hompel, 2020. Der Industrie 4.0 Maturity Index in der betrieblichen Anwendung – aktuelle Herausforderungen, Fallbeispiele und Entwicklungstrends (acatech Kooperation), München.

[5] Bose, R.J.C., Mans, R.S., van der Aalst, W.M. Wanna improve process mining results?

[6] Galic, G., Wolf, M., Salzmann, O., Unger, T., 2021. Delivering Value with Process Analytics: Process Mining adoption and success factors.

[7] Oliveira, P., Rodrigues, F., Henriques, P.R., 2005. A Formal Definition of DQ Problems.doc.

[8] Flack, C., Dreher, S., Birk, A., Wilhelm, Y., 2020. Process Mining in der Produktion: Spezifische Herausforderungen bei der Anwendung 115 (11).

[9] Reinkemeyer, L., 2020. How to Get Started, in: Reinkemeyer, L. (Ed.), Process Mining in Action. Springer International Publishing, Cham.

[10] Jahn, M., 2017. Industrie 4.0 konkret. Springer Fachmedien Wiesbaden, Wiesbaden.

[11] El Alaoui, I., Gahi, Y., Messoussi, R., 2019. Big Data Quality Metrics for Sentiment Analysis Approaches, in: Proceedings of the 2019 International Conference on Big Data Engineering. BDE 2019: 2019 International Conference on Big Data Engineering, Hong Kong Hong Kong. 11 06 2019 13 06 2019. Association for Computing Machinery, New York,NY,United States.

[12] Mocnik, F.-B., Zipf, A., Fan, H., 2017. Data Quality and Fitness for Purpose.

[13] Bitkom-Gremium - AK Big Data, 2015. Leitlinien für den Big-Data-Einsatz. https://www.bitkom.org/sites/default/files/file/import/150901-Bitkom-Positionspapier-Big-Data-Leitlinien.pdf. Accessed 13 November 2021.

[14] Wank, A., 2019. Methodik zur Wertstromintegration einer aktiven Bauteilrückverfolgung in die diskrete Variantenfertigung. Shaker, Herzogenrath.

[15] Rajpurohit, A., 2013. Big data for business managers — Bridging the gap between potential and value, in: 2013 IEEE International Conference on Big Data. Silicon Valley, California, USA, 6 - 9 October 2013. 2013 IEEE International Conference on Big Data, Silicon Valley, CA, USA. 06.10.2013 - 09.10.2013. IEEE, Piscataway, NJ.

[16] ZVEI. ZVEI-Traceability-Initiative "Traceability-Levels für Produktkategorien".

[17] Batini, C., Cappiello, C., Francalanci, C., Maurino, A., 2009. Methodologies for data quality assessment and improvement. ACM Comput. Surv. 41 (3).

[18] Urnauer, C., Schreiber, M., Bausch, P., Metternich, J., 2021. Anwendungen aktiver Traceability-Systeme: Datennutzung in der digitalisierten Produktion. ZWF - Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (3).

[19] Ceravolo, P., Azzini, A., Angelini, M., Catarci, T., Cudré-Mauroux, P., Damiani, E., Mazak, A., van Keulen, M., Jarrar, M., Santucci, G., Sattler, K.-U., Scannapieco, M., Wimmer, M., Wrembel, R., Zaraket, F., 2018. Big Data Semantics.

[20] Schreiber, M., Bausch, Phillip, Best, Julian, Metternich, J., 2020. Datenanalyse in Produktionsprozessen: Potenziale und Herausforderungen des Process-Mining-Einsatzes in Theorie und betrieblicher Praxis. ZWF - Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (5).

[21] Ehrlinger, L., Rusz, E., Wöß, W., 2019. A Survey of Data Quality Measurement and Monitoring Tools.

[22] Rahm, E., Do, H.H., 2000. Data Cleaning: Problems and Current Approaches.

[23] Karkouch, A., Mousannif, H., Al Moatassime, H., Noel, T., 2016. Data quality in internet of things: A state-of-the-art survey. Journal of Network and Computer Applications 73.

[24] Laranjeiro, N., Soydemir, S.N., Bernardino, J., 2015. A Survey on Data Quality: Classifying Poor Data.

[25] Ge, M., Helfert, M. A Review Of Information Quality Research.

[26] Ranjit Singh, Dr. Kawaljeet Singh, 2010. A Descriptive Classification of Causes of Data Quality Problems in Data Warehousing.

[27] Strong, D.M., Lee, Y.W., Wang, R.Y., 1997. Data quality in context. Commun. ACM 40 (5).

[28] Verhulst, 2016. Evaluating quality of event data within event logs an extensible framework.

[29] Wang, R.Y., Reddy, M.P., Kon, H.B., 1995. Toward quality data: An attribute-based approach. Decision Support Systems 13 (3-4).

[30] Wang, R.Y., Strong, D.M., 1996. Beyond Accuracy: What Data Quality Means to Data Consumers. Journal of Management Information Systems 12 (4).

[31] Woodall, P., Oberhofer, M., Borek, A., 2014. A classification of data quality assessment and improvement methods. IJIQ 3 (4).

Biography

Markus Schreiber, M. Sc. has been a researcher and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2018. His research interests include the traceability of components and equipment, as well as the optimization of production processes.

Prof. Dr.-Ing. Joachim Metternich has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2012. In addition, Prof. Metternich is the spokesman for the Center of Excellence for Medium-Sized Businesses in Darmstadt and president of the International Association of Learning Factories.

3rd Conference on Production Systems and Logistics

A Circular Economy Strategy Selection Approach: Component-based Strategy Assignment Using the Example of Electric Motors

Martin Benfer[1], Patrizia Gartner[1], Felix Klenk[1], Christoph Wallner[2],
Marie-Christin Jaspers[1], Sina Peukert[1], Gisela Lanza[1]
[1] *Karlsruhe Institute of Technology (KIT), wbk Institute of Production Science, Kaiserstr. 12, 76131 Karlsruhe, Germany*
[2] *Loyola Marymount University (LMU), Seaver College of Science and Engineering, 1 LMU Dr, Los Angeles, CA 90045, USA*

Abstract

The sustainability of industrial processes and products is a core issue of our time. There are several approaches to move from a linear, inherently wasteful economic principle to a circular economy focused on conserving products, resources, and energy. However, selecting which of the circular economy strategies ranging from reuse, repurpose, and remanufacture to recycling is crucial to ensure the economic viability of the product. This contribution proposes an iterative, component-based circular economy strategy selection method that supports product and production planners in choosing the appropriate circular economy strategy. For this approach, the suitability of each component for circular economy strategies is assessed based on identified key properties. In case of no fitting strategy, further component decomposition is devised, and the process is repeated. To further support the design of circular economy strategies, a modular process build set is suggested, enabling the swift composition of the processing sequence. The approach is then applied to the example of an electric motor of a battery electric vehicle. The presented approach allows a quick first assessment of the viability of different circular economy strategies and helps product and production engineers develop product-specific circular economy strategies.

Keywords

Circular Economy; Strategy Selection; Sustainable Production; Electromobility; Product-Process Co-Design;

1. Introduction

Due to rising environmental concerns, the issue of sustainability and the preservation of resources is increasingly seen as a fundamental challenge of industrial production. The classical, so-called linear economy model inherently consumes finite resources and creates non-usable waste. The circular economy (CE) approach conserves the value of products that are no longer in use by restoring their functionality, using them for different purposes, or transforming them into new products. The available circular economy strategies have been described extensively, most comprehensively by the 9R model [1]. Each strategy preserves different aspects of a product, so that the question of which strategy to choose arises. The consensus is to prefer the closest loop possible, thereby retaining as much value inherent in a product and performing the least work [2]. However, deciding what exactly denotes what is possible is difficult, especially since many factors have to be considered, including economic feasibility. This issue is especially prominent when creating the structures to enable a circular economy for an existing or new product. In this case, several different options have to be weighed to find the best solution. This problem is complicated by

DOI: https://doi.org/10.15488/12133 publish-Ing.

the complexity of many modern products, especially if individual components are best suited for different CE strategies. While it is technically possible to calculate business cases for each combination of CE strategies, this quickly becomes prohibitively laborious. This paper thus proposes a methodology to determine fitting CE strategies based on an assessment of the properties of the examined product. It further offers an approach to define the necessary industrial processes to facilitate the chosen CE strategies. This overall approach is applied in an exemplary case of an electric car drive, showcasing its potential.

This paper is structured as follows: first, section 2 examines the fundamentals of circular economy and summarizes the current state of research regarding the issue described above. Next, section 3 introduces the methodology for selecting fitting CE strategies and creating matching processes. This approach is then applied exemplarily on an electric drive. Subsequently, section 5 discusses the potential applications and limitations of this approach. Finally, section 6 offers a summary and outlook.

2. Fundamentals of Circular Economy

The extraction and processing of natural raw materials are the cornerstones of today's industrial production. The decreasing cost and increasing efficiency in their extraction and processing have been a basis for the growth of the world economy and the improvement in global living standards for years [3]. Thus, today's manufacturing industry mainly relies on a linear economy model characterized by a "take-make-use-dispose" mindset [2]. To preserve the biodiversity and habitable characteristics of the planet, there are good industrial and societal reasons to increase the use of secondary resources in manufacturing, i.e., through the recovery of materials and resources from end-of-life products [3]. According to a study from the *Circular Economy Initiative Germany*, a complete transition to a circular economy can reduce consumption of natural resources by 50 % until 2050 compared to 2018 [4]. A circular economy is thereby defined as an industrial system, which is restorative and regenerative by design [2], and by means of which products, components, and materials are kept at their highest utility and value along the entire life cycle [5]. In literature, many frameworks already capture the circular economy, of which the 9R model is the most comprehensive [1]. These 9R describe different CE strategies, which include: (1) *refuse*, i.e., preventing the use of raw materials in the first place, through (2) *reduce* and (3) *reuse*, product recovery options like (4) *repair*, (5) *refurbish* and (6) *remanufacture*, to (7) *repurpose*, and lastly (8) *recycle* and (9) *energy recovery* [1]. With every increasing step of the 9R model, the level of circularity, i.e., the volume of raw material extraction and negative environmental impact decrease. A high level of circularity improves economic, social, and ecological value creation. However, feasibility and different characteristics, e.g., product composition, market and competitor situation, or legal and governmental restrictions, need to be considered [2,1].

This paper is based on the 9R model. However, some alterations are made to accommodate the application in the design of circular value streams. Thus, the options refuse and reduce focussed on product design and resource procurement are foregone. The options repair, refurbish and remanufacture are considered jointly under the term remanufacture, as they are mostly distinguished by the degree of alteration necessary to restore functionality. Energy recovery is considered a non-desirable outcome of the strategy selection but may be necessary in some instances. Accordingly, the CE-strategies considered here are: *reuse*, the use of a functional product in a similar application without significant alterations, *repurpose*, the use of a largely functional product in a different application, *remanufacture*, the restoration of a non-functional product while preserving existing functionality, and *recycle*, the destructive utilization of materials in a product for new products. [6]

While a CE strategy describes the general idea and recovered aspects of a product, the particular CE process sequence still needs to be developed afterward. Unfortunately, there are no universally applicable process chains for CE strategies. On the contrary, many remanufacturing processes, for example, are designed differently and depend on the material, product layout, and different product or market-related criteria [7].

Typical processes for remanufacturing are described by [8], while [9] describe recycling processes. However, while these process chains are different, they still consist of common elementary processes. According to [10], those are the collection and sorting of the waste, the inspection of the individual parts, and finally, the disassembly.

3. CE Strategy Evaluation, Selection, and Process Development

As described above, selecting an appropriate CE strategy for a given product is crucial for achieving sustainability and economic viability. Thus, other methods for strategy evaluation and selection are necessary. There are several comprehensive but complex methods to determine suitable CE strategies. [11] propose a comprehensive multi-criteria decision tool to select suitable CE strategies. [12] assign CE strategies to product instances using Bayesian updating and fuzzy set theory at their EOL. [13] use an analytical hierarchy process and case-based reasoning to determine CE strategies based on similarity and consider both a product and a component level. Several other approaches exist, that examine the feasibility of remanufacturing products [14,15]. [16] present a tool for evaluation of product recycling using the concept of information entropy. Other approaches implement the preference for smaller circles by using a cascade model that prioritizes closer cycles wherever possible [17]. In terms of process development, most contributions are focused on describing relevant processes. [5] highlight the most important system level problems, methods and tools for re- and demanufacturing within the CE context. [18] discuss typical remanufacturing process steps in detail and provide a process sequencing model for cost minimization. [19] determine an optimal remanufacturing process sequences depending on product conditions. While several approaches exist in literature, most are focused on the most comprehensive evaluation. Furthermore, most approaches only consider the selection on the product level, instead of considering different options for each component. Finally, typically, the selection of strategies and their detailing is not considered jointly.

4. Methodology for CE Strategy Selection and Process Development

The selection and development of suitable CE strategies and processes require a comprehensive assessment of the examined product and its components. Furthermore, different options for the CE processes should be considered and assessed before a fitting solution is finalized. Thus, this paper proposes a three-stage approach, based on the selection of relevant CE strategies from the 9R model described in section 2. The approach is shown in Figure 1. The overall process presented here was developed considering the following principles: 1) Use as little information as possible in every step to limit the necessary effort, 2) Generalize concepts where possible to aid applicability in different contexts, and 3) Detail as late as possible.

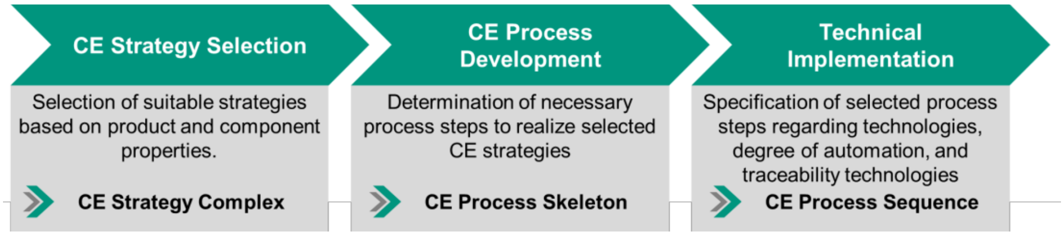


Figure 1: Proposed Three-Stage Process for CE Strategy Selection, Process Development, and Implementation

In the first stage, the fundamental properties of the examined product and its components are assessed. Then, based on this assessment, the suitability of CE strategies for the components is evaluated and a fitting CE strategy complex comprising CE strategies for all components is established. Subsequently, the selected strategies are detailed using a building set of CE processes in stage two. Finally, process variants and automation potentials are considered in stage three, and necessary information exchange technologies are laid out. After this process, a comprehensive concept for designing a circular value chain for a specific product exists. The required steps are specified in the subsequent sections.

4.1 CE Strategy Selection

The first step of the process is the selection of fitting CE strategies for products and components. To assess whether a CE strategy is suitable for a product or component, first, the relevant characteristics of each CE strategy have to be considered. These can then be matched with product characteristics to obtain a suitability score for each product-strategy combination. For the assessment products and components are considered in their average end-of-life (EOL) state. To determine the criteria that describe the suitability of a strategy, first, a qualitative literature study of CE strategy characteristics was performed. With these first results, expert interviews were conducted to weigh the relevance of different criteria using pairwise analysis. The experts were selected from both academia and industry and covered product, production, and circular economy knowledge. Table 1 shows the selected criteria and their respective weights for each CE strategy. In addition to four considered CE strategies, the disassembly of the product or component is also evaluated as an option.

Table 1: CE Strategy Selection Criteria and Respective Weights for each CE Strategy based on Pairwise Comparison

Criteria	**Description**	**Reuse**	**Repurpose**	**Remanufacture**	**Recycle**	**Disassembly**
EOL Condition	Overall condition of the component at EOL (wear, damage, defects)	0.45	0.38	0.32		0.15
Technical Relevance	Component still usable in its design form, corresponds to the state of the art	0.30		0.14		0.15
Volume	Demand, quantity of products	0.08		0.09	0.08	
Residual Value	Age, material value of the unit	0.08	0.08	0.09		
Applicability	Usability in other applications		0.38			
Conversion Effort	Cost and effort for use in other applications		0.08			
Remanufacturing Effort	Cost and effort for reconditioning the parts to new condition			0.27		
Raw Material Value	Value of the included materials/raw materials				0.45	
Disassembly Effort	Cost and effort of recovering the raw materials/materials				0.23	0.15
Recycling Rate	Recyclable portion / total part				0.15	
Modularity	Possibility of disassembly into functional subunits					0.45
Cycle Preference	Preference for CE strategy	4	3	2	0	1

To perform the selection process each of the criteria need to be evaluated for an examined product, for example using a 0 to 4 scale. Then the overall score $v_{p,s}$ for each product p and strategy s is expressed as

$$v_{p,s} = \sum_{i \in C} c_{p,i} w_{s,i} + 0.1 cp_s \qquad (1)$$

Where $c_{p,i}$ is the product specific criteria score, $w_{s,i}$ the strategy specific criteria weight and cp_s is the cycle preference of the strategy. The cycle preference is added to create a preference for smaller cycles as described by [2]. Subsequently, the highest-scoring strategy is selected. If disassembly is selected, the most natural decomposition of the product into components has to be determined. For each of those components, the process is then repeated. If multiple best scoring strategies are within a close range, each option should be considered in more detail. The result of the evaluation model is summarized in a strategy complex, as shown in Figure 2. The model provides a general estimate of the practical utility of all considered CE strategies. The process complex resulting from the evaluation model indicates which CE strategies are targeted for the respective component and its sub-components.

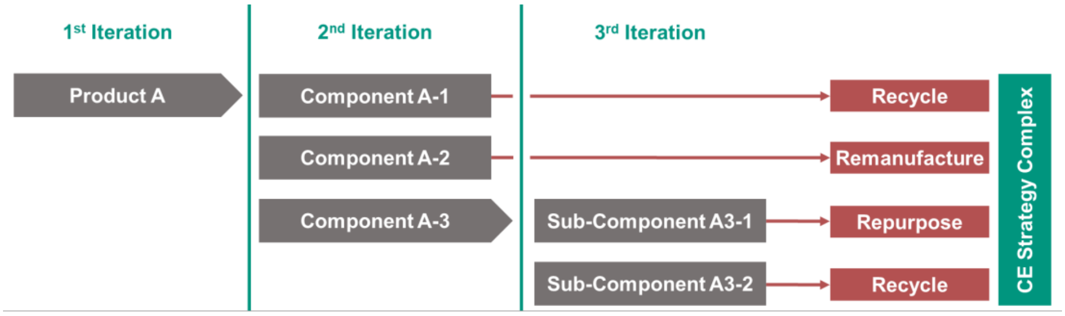
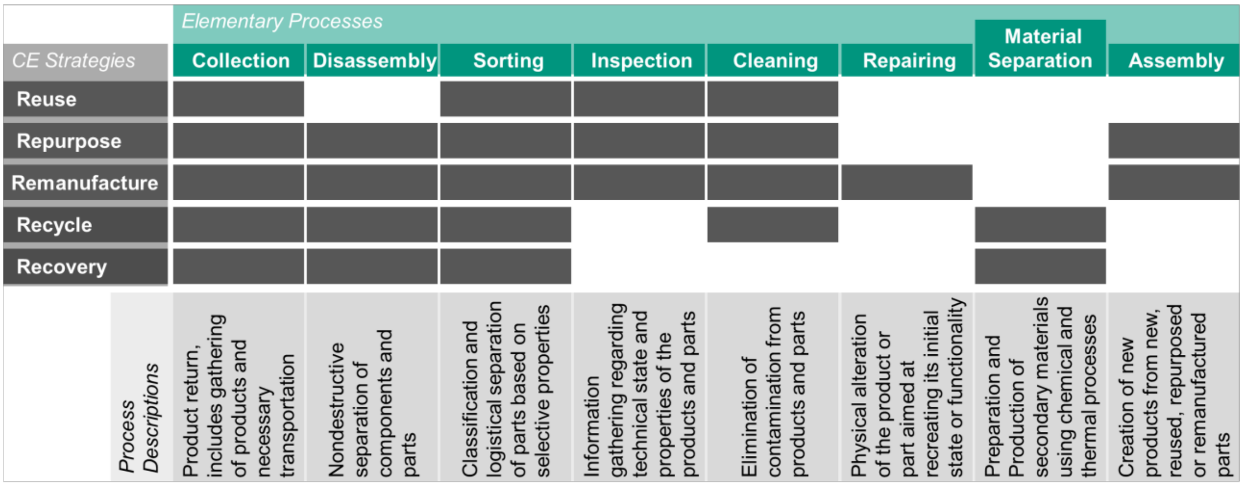


Figure 2: Iterative Selection of Strategies and Resulting Strategy Complex

4.2 CE Process Development

After establishing a CE strategy complex for a given product, a process sequence is developed. As noted in section 2, CE strategies consist of several processes depending on the examined product. It is noticeable that the CE strategies share similar processes and differ mainly in the arrangement, number, or combination of those. Thus, this contribution proposes a modular process build set, containing multiple elementary processes, that enables the swift composition of process sequences. As indicated in section 2, different CE strategies share several elementary processes which can be utilized in the process build set. The considered elementary processes and their relevance for the different CE strategies are shown in Figure 3. The elementary processes are *collection*, *disassembly*, *sorting*, and *inspection*, as proposed by [10]. Additionally, *cleaning*, *repairing*, *material separation*, and *assembly* are included. These elementary processes where identified through a qualitative commonality analysis in CE process descriptions in scientific literature. Even though some other processes exist, the presented elementary processes represent the core of the typical CE strategies. The CE strategy recovery is included in the build set, as it may be necessary for some defective product instances.



	Elementary Processes							
CE Strategies	**Collection**	**Disassembly**	**Sorting**	**Inspection**	**Cleaning**	**Repairing**	**Material Separation**	**Assembly**
Reuse	■		■	■	■			
Repurpose	■	■	■	■	■			■
Remanufacture	■	■	■	■	■	■		■
Recycle	■	■	■		■		■	
Recovery	■	■	■				■	
Process Descriptions	Product return, includes gathering of products and necessary transportation	Nondestructive separation of components and parts	Classification and logistical separation of parts based on selective properties	Information gathering regarding technical state and properties of the products and parts	Elimination of contamination from products and parts	Physical alteration of the product or part aimed at recreating its initial state or functionality	Preparation and Production of secondary materials using chemical and thermal processes	Creation of new products from new, reused, repurposed or remanufactured parts

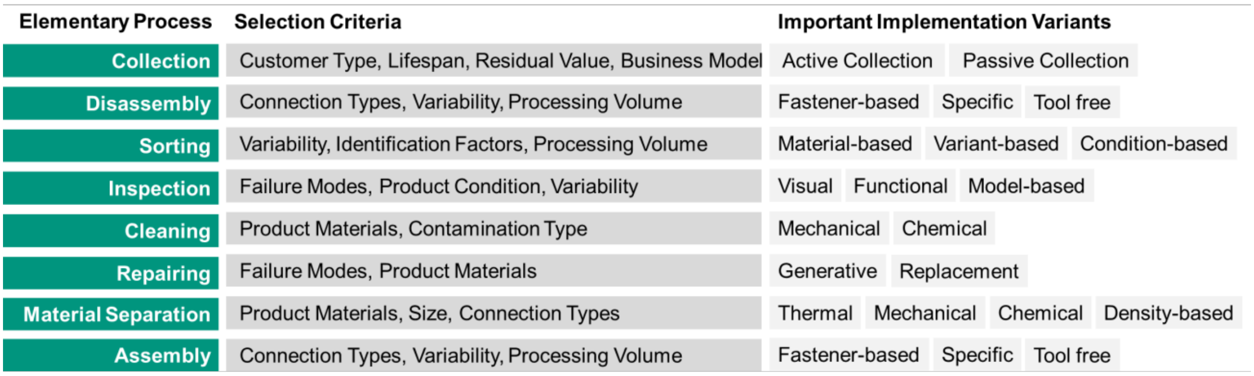
Figure 3: Elementary Processes and their Application in Different CE Strategies

A process skeleton is devised using the strategy complex developed in 4.1 and the relevant elementary processes for each strategy. The resulting process sequence is modelled using the event-driven process chain (EPC). As the previously selected CE strategies only specify the process for an assumed general product instance, an inspection, sorting, and deviation handling needs to be implemented before every CE process.

The deviation handling allows for the recovery of specific product instances not corresponding with CE strategy requirements.

4.3 Technological Implementation

Next, the elementary processes need to be specified for use with the particular product or component for the technological implementation. For each elementary process, criteria to help define the suitable process implementation variant were identified. The implementation variants were determined by analysing CE process descriptions and identifying commonly described variants. Based on the recognised variants, influencing criteria on the selection of process variants were determined. Depending on those criteria, a suitable process variant can be selected. As this criteria and variants are significantly more complex compared to the decisions made in 4.1, no deterministic selection scheme is proposed. Instead, the specific criteria serve as the basis for a more comprehensive decision process regarding the variants, utilizing expert knowledge. The relevant criteria for each elementary process and important process variants are shown in Figure 4.



Elementary Process	Selection Criteria	Important Implementation Variants			
Collection	Customer Type, Lifespan, Residual Value, Business Model	Active Collection	Passive Collection		
Disassembly	Connection Types, Variability, Processing Volume	Fastener-based	Specific	Tool free	
Sorting	Variability, Identification Factors, Processing Volume	Material-based	Variant-based	Condition-based	
Inspection	Failure Modes, Product Condition, Variability	Visual	Functional	Model-based	
Cleaning	Product Materials, Contamination Type	Mechanical	Chemical		
Repairing	Failure Modes, Product Materials	Generative	Replacement		
Material Separation	Product Materials, Size, Connection Types	Thermal	Mechanical	Chemical	Density-based
Assembly	Connection Types, Variability, Processing Volume	Fastener-based	Specific	Tool free	

Figure 4: Elementary Processes, Relevant Selection Criteria, and Important Implementation Variants,

Furthermore, each process can be automated to varying degrees, similar to different degrees of automation in linear production. To determine the fitting degree of automation, the variability or the condition of the processed products and components also need to be considered. Historically, the processes of disassembly, inspection, and repair have been challenging to automate, as they are highly dependent on product condition. However, more recently, traceability technologies have emerged that aid product and component identification and facilitate automating inspection by using condition monitoring [20]. To enable this automation in CE processes, implementing a specific traceability system may be beneficial [21]. Also, significant improvements have been made in terms of adaptive robot-based disassembly and repair systems.

With the conclusion of the technological implementation, the first draft of CE processes for an existing or currently developed product is derived. An example of such a process draft is shown in Figure 6. This draft may be used to gauge the viability of CE strategies for a particular product and can serve as a basis for more detailed planning of the overall CE strategy and the necessary processes.

5. Validation

Battery electric vehicles (BEV) have seen rapid growth in the last decade. BEV's are intended to replace cars with internal combustion engines, reducing personal mobility's environmental impact. As batteries and many electric drives rely on specific, naturally limited resources, designing fitting CE strategies is vital for sustainability. For this paper, the CE strategies for a permanent magnet synchronous motor (PMSM) are selected and detailed. In this case, it is assumed that the majority of the PMSM is still functional at the EOL of the vehicle. The design of PMSM for automotive appliances follows a typical structure consisting of a *casing*, an *external stator*, and an *internal rotor*. The casing is typically manufactured as one-piece casting and is sealed by an endplate. The stator contains copper coil windings with different configurations depending on the winding method. The rotor consists of a solid rotor shaft on which the stacked plates are joined and which is mounted in the housing with bearings. The permanent magnets are attached to the stacked plates by pressing or gluing. [22]

The three-step model introduced in section 3 is applied to the PMSM and its components and sub-components. The scoring of the evaluation criteria is established, and the result of each iteration is summarized. The scoring of the individual evaluation criteria was based on information from expert interviews. In the *1st Iteration,* the full EOL PMSM was examined. PMSM in their EOL state are still functional. However, various minor damages can occur to the motor. Due to the contained valuable raw materials such as copper, the raw material value and the recycling rate of the motor are high. Moreover, the overall residual value is high due to the overall good condition and the extensive value-adding processes in the manufacturing process. Although further development of electric drives is expected, technical relevance compared to then low-budget PMSM is conceivable. Due to the design, dismantling to the next smaller sub-components is relatively easy. The result of the first iteration was thus further disassembly. For the *2nd Iteration*, the sub-components casing, stator, and rotor are considered. The casing is made up of predominantly homogeneous materials. Here the residual value of the materials and the good EOL state are dominant. At the end of the evaluation, the casing is eligible for remanufacturing or recycling. The rotor has high residual and material value due to the contained copper, which is easily recyclable. Thus, recycling is selected. The rotor of a PMSM consists of a solid steel rotor shaft that carries the laminated cores with the permanent magnets. The condition and residual value of the rotor can be rated as good. At the same time, the material value of the installed components varies greatly. However, the sub-components can easily be further dismantled. Therefore, the result of the evaluation initiates another iteration. The *3rd Iteration* is conducted for the sub-components of the rotor. These are identified as permanent magnets and rotor shafts with laminated cores. The overall result of the strategy selection is shown in Figure 5.

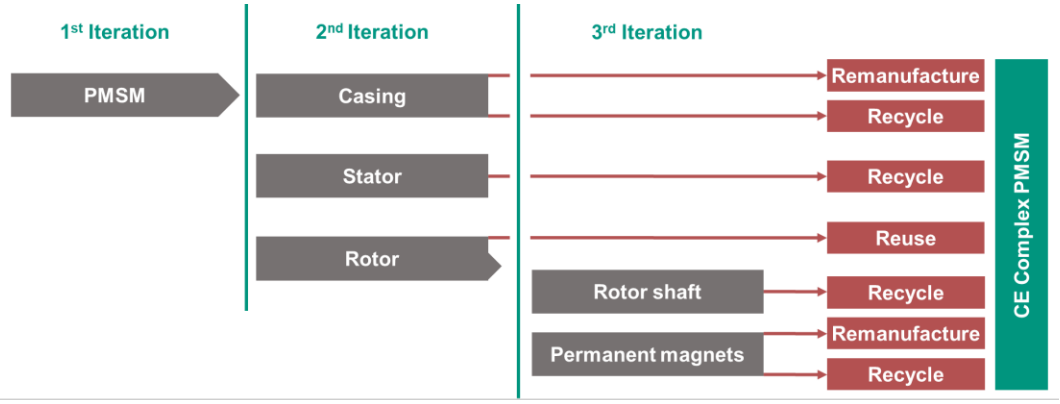


Figure 5: Resulting Strategy Complex for PMSM

The next step was the definition of a more detailed process chain. For this purpose, first the disassembly, inspection, and sorting sequence was connected with each of the strategies. Then each strategy was detailed using the process build set introduced in 4.2. In case the previous step selected multiple potential strategies, the strategy with the highest circle preference was used in the process development. Figure 6 shows the resulting process chain.

To arrive at a detailed process draft, each process step depicted in Figure 6 was further detailed using the methodology discussed in section 4.3. This included the selection of process variants based on product properties. Furthermore, fitting automation levels were selected. To enable a cost-efficient circular economy, a high level of automation is desirable. To facilitate a high degree of automation, a traceability system was planned that allows automated inspection and sorting by offering information on the likely condition of rotor and casing based on usage sensors and identification of product variants. The resulting process chain can be used to plan a circular production system for PMSM. It shows how a first draft of this system can be created with relatively little effort. This draft helps to evaluate whether existing process chains should be reconsidered and guides more detailed considerations of the different available options.

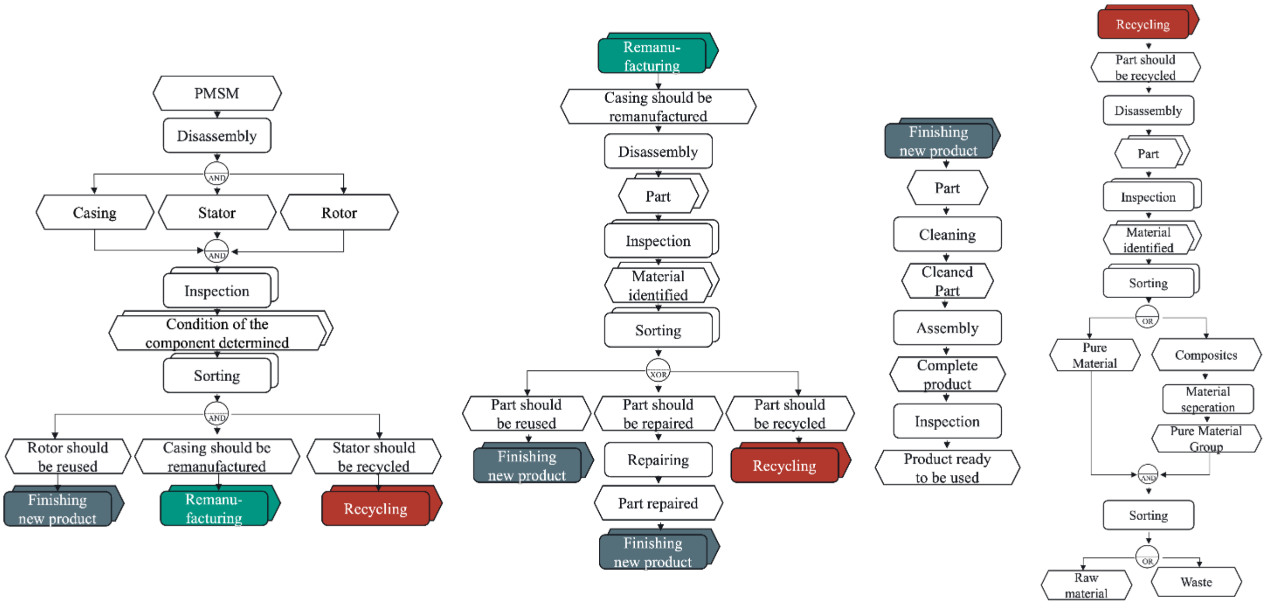


Figure 6: Determined Process Chain for PMSM, Colours Indicate Different Strains of the Process Chain

6. Discussion

Although the approach presented in this paper aims at a comprehensive decision support method for circular economy strategy selection, its utilization and advantageousness are subject to some restrictions. First, the application is most suitable for the system design phase of a circular economy. It assumes many degrees of freedom, i.e., regarding the product design, market and competitor situation, or design of business models and closed-loop supply chains. Moreover, a sufficient return of EOL products in terms of volume and quality is assumed to be profitable. Second, the proposed model only partially considers the organization's specifics regarding its overall business model, capabilities, and influencing factors. Thus, the model may be used to develop an initial CE strategy and process complex that can be refined using more detailed planning procedures. Lastly, the approach relies on subjective assessments of the decision-makers. Although the subjectivity can be reduced, e.g., through comparative analysis and inclusion of multiple perspectives, there remains at least some level of subjectivity, leading to different outcomes for different decision makers.

7. Summary and Outlook

This contribution proposed a methodology for selecting and configuring CE strategies in an early design stage. The methodology combines a comprehensive consideration of product properties with high practicality and may thus limit the effort necessary to derive sensible CE strategies. As many organizations today face the challenge of quickly adapting to a CE paradigm, the proposed method could help focus and guideline the transformation.

In the future, the methodology could be expanded by including a more quantitative analysis regarding the economic and ecological viability of the different options. Additionally, the effect of other criteria on strategy suitability, estimated based on expert interviews here, could be examined empirically. Further research on cross-company traceability technologies is necessary as they have shown significant potential in enabling the automation of CE processes. Finally, a further investigation of product design for CE strategies is essential to enable more effective and efficient CE processes while retaining product performance.

Acknowledgments

This research was funded by the Ministry of Economic Affairs, Labour and Tourism of Baden-Württemberg (Germany) as part of the research project PoTracE and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 432466774.

References

[1] van Buren, N., Demmers, M., van der Heijden, R., Witlox, F., 2016. Towards a Circular Economy: The Role of Dutch Logistics Industries and Governments. Sustainability 8 (7), 647.

[2] Ellen MacArthur Foundation, 2013. Towards the circular economy Vol. 1: an economic and business rationale for an accelerated transition. https://www.ellenmacarthurfoundation.org/publications/towards-the-circular-economy-vol-1-an-economic-and-business-rationale-for-an-accelerated-transition. Accessed 2 February 2022.

[3] OECD, 2019. Global Material Resources Outlook to 2060: Economic Drivers and Environmental Consequences, Paris. https://www.oecd.org/env/global-material-resources-outlook-to-2060-9789264307452-en.htm. Accessed 2 February 2022.

[4] Kadner, S., Kobus, J., Hansen, E.G., Akinci, S., Elsner, P., Hagelüken, C., Jaeger-Erben, M., Kick, M., Kwade, A., Müller-Kirschbaum, T., Kühl, C., Obeth, D., Schweitzer, K., Stuchtey, M., Vahle, T., Weber, T., Wiedemann, P., Wilts, H., Wittken, R. von, 2021. Circular Economy Roadmap for Germany. https://www.acatech.de/publikation/circular-economy-roadmap-fuer-deutschland/. Accessed 2 February 2022.

[5] Tolio, T., Bernard, A., Colledani, M., Kara, S., Seliger, G., Duflou, J., Battaia, O., Takata, S., 2017. Design, management and control of demanufacturing and remanufacturing systems. CIRP Annals 66 (2), 585–609.

[6] Potting, J., Hekkert, M.P., Worrell, E., Hanemaaijer, A. Circular economy: measuring innovation in the product chain. https://dspace.library.uu.nl/handle/1874/358310. Accessed 2 February 2022.

[7] Niemann, J., Schuh, G., Baessler, E., Eigner, M., Stolz, M., Steinhilper, R., Janusz-Renault, G., Hieber, M., 2009. Management des Produktlebenslaufs, in: Bullinger, H.-J., Spath, D., Warnecke, H.-J., Westkämper, E. (Eds.), Handbuch Unternehmensorganisation. Strategien, Planung, Umsetzung, 3., neu bearbeitete Auflage ed. Springer, Berlin, Heidelberg, pp. 223–315.

[8] Lee, C.-M., Woo, W.-S., Roh, Y.-H., 2017. Remanufacturing: Trends and issues. Int. J. of Precis. Eng. and Manuf.-Green Tech. 4 (1), 113–125.

[9] Ortegon, K., Nies, L., Sutherland, J.W., 2014. Recycling, in: Laperrière, L., Reinhart, G., Chatti, S., Tolio, T. (Eds.), CIRP encyclopedia of production engineering. With 85 tables. Springer, Berlin, pp. 1039–1042.

[10] Sundin, E., Elo, K., Mien Lee, H., 2012. Design for automatic end-of-life processes. Assembly Automation 32 (4), 389–398.

[11] Alamerew, Y.A., Brissaud, D., 2019. Circular economy assessment tool for end of life product recovery strategies. Jnl Remanufactur 9 (3), 169–185.

[12] Pochampally, K.K., Vadde, S., Kamarthi, S.V., Gupta, S.M., 2004. Beyond sensor-assisted diagnosis of used products, in: Environmentally Conscious Manufacturing IV. Optics East, Philadelphia, PA. Monday 25 October 2004. SPIE, pp. 138–146.

[13] Ghazalli, Z., Murata, A., 2011. Development of an AHP–CBR evaluation system for remanufacturing: end-of-life selection strategy. International Journal of Sustainable Engineering 4 (1), 2–15.

[14] Goodall, P., Rosamond, E., Harding, J., 2014. A review of the state of the art in tools and techniques used to evaluate remanufacturing feasibility. Journal of Cleaner Production 81, 1–15.

[15] Rizova, M.I., Wong, T.C., Ijomah, W., 2020. A systematic review of decision-making in remanufacturing. Computers & Industrial Engineering 147, 106681.

[16] Bognar, N., Rickert, J., Mennenga, M., Cerdas, F., Herrmann, C., 2019. Evaluation of the Recyclability of Traction Batteries Using the Concept of Information Theory Entropy, in: Pehlken, A. (Ed.), Cascade Use in Technologies 2018. Internationale Konferenz Zur Kaskadennutzung und Kreislaufwirtschaft - Oldenburg 2018. Vieweg, Berlin, Heidelberg, pp. 93–103.

[17] Kalverkamp, M., Pehlken, A., Wuest, T., 2017. Cascade Use and the Management of Product Lifecycles. Sustainability 9 (9), 1540.

[18]Li, J., Wu, Z., 2014. Remanufacturing Processes, Planning and Control, in: , New Frontiers of Multidisciplinary Research in STEAM-H (Science, Technology, Engineering, Agriculture, Mathematics, and Health). Springer, Cham, pp. 329–356.

[19]Kin, S.T.M., Ong, S.K., Nee, A., 2014. Remanufacturing Process Planning. Procedia CIRP 15, 189–194.

[20]Gartner, P., Benfer, M., Kuhnle, A., Lanza, G., 2021. Potentials of Traceability Systems - a Cross-Industry Perspective. Procedia CIRP 104, 987–992.

[21]Benfer, M., Gartner, P., Treber, S., Kuhnle, A., Häfner, B., Lanza, G., 2020. Implementierung von unternehmensübergreifender Traceability. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (5), 304–308.

[22]Röth, T., Kampker, A., Deutskens, C., Kreisköther, K., Heimes, H.H., Schittny, B., Ivanescu, S., Büning, M.K., Reinders, C., Wessel, S., Haunreiter, A., Reisgen, U., Thiele, R., Hameyer, K., Doncker, R.W. de, Sauer, U., van Hoek, H., Hübner, M., Hennen, M., Stolze, T., Vetter, A., Hagedorn, J., Müller, D., Rewitz, K., Wesseling, M., Flieger, B., 2018. Entwicklung von elektrofahrzeugspezifischen Systemen, in: Kampker, A., Vallée, D., Schnettler, A. (Eds.), Elektromobilität. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 279–386.

Biography

Martin Benfer (*1994) is a research associate and doctoral student at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

Patrizia Gartner (*1993) is a research associate and doctoral student at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

Felix Klenk (*1993) is a research associate and doctoral student at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

Christoph Wallner (*1997) is a former Bachelors's Student at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT). He is currently a Masters's Student at Seaver College of Science and Engineering at Loyola Marymount University (LMU), Los Angeles.

Marie-Christin Jaspers (*1996) is a Masters's Student at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

Sina Peukert (*1991) is Group Leader for Global Production Strategies at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

Gisela Lanza (*1973) is the Director of Production Systems Division at the wbk Institute of Production Science at Karlsruhe Institute of Technology (KIT).

3rd Conference on Production Systems and Logistics

Practical Requirements For Digital Twins In Production And Logistics

Hendrik van der Valk[1], Jan-Luca Henning[1], Stephanie Winkelmann[1], Hendrik Haße[2]

[1] *Chair for Industrial Information Management, TU Dortmund University, Dortmund, Germany*
[2] *Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany*

Abstract

Companies are under tremendous pressure to analyze and optimize their productional and logistical networks in today's global business world. Hence, practitioners and researchers show great interest in digital twins. A digital twin is a virtual construct that mirrors real-world objects and conceptual ideas while it processes, handles, distributes, and optimizes data streams. Its main purpose is to visualize, analyze, and optimize objects and systems, making a digital twin highly suitable to help companies gain an advantage over their competitors through a great degree of transparency over their production and logistics. Therefore, almost every company evaluates the usage of digital twins. Nevertheless, many companies struggle to instantiate digital twins since they lack fundamental knowledge about all necessary components of a digital twin and the individual requirements for the operation of the digital twin. This lack of knowledge hinders the broad implementation in practice. Research shows many descriptions of theoretical use cases and field studies but rarely describes digital twins in real operational settings. To address this research gap between theoretical concepts and practical challenges of the implementation of digital twins, this paper investigates the practical requirements of digital twins in real-life usage. Based on a thorough interview series with international manufacturing and logistics experts, we identify and analyze the requirements for data handling, data policy, and services of digital twins and cluster them according to the requirements engineering approach. Through a comprehensive overview of the different requirements, the paper delivers profound insights into the needs of companies from various fields and, therefore, gives practitioners a guideline on crucial aspects of implementing digital twins.

Keywords

Digital Twins; Production; Logistics; Practical Requirements; Interview Study

1. Introduction

Modern global business environments face significant pressure to perform in a holistic analyzed, and optimized setting. This mainly includes production as well as logistics networks. Hence, a thorough search for digital tools helping with the optimization processes has started. A promising tool for these tasks is the digital twin. Thus, practitioners and researchers are interested in digital twins [1,2]. A digital twin mirrors physical counterparts, processes data, and provides the opportunity to create optimized processes [3]. Significantly, the digital twin`s primary purposes, including visualization, analysis, and optimization of objects and systems, provide appropriate resources for companies to gain an advantage over their competitors by providing a high level of visibility into their production and logistics [4]. As a result, almost every company is evaluating the use of digital twins [5,6].

DOI: https://doi.org/10.15488/12178

However, due to the significant gap between scientific research and practical knowledge, companies still struggle to implement digital twins suitable for real-world applications. Of considerable relevance are the topics of the necessary components of a digital twin and the individual requirements for operating the digital twin. This lack of knowledge blocks widespread implementation in practice leading to the following research questions:

RQ1: What are the individual requirements for digital twins in production and logistics?

To bridge the gap between scientific knowledge and practical implementation, users need to know the individual requirements. To gain an industry-focused result, we aim at a practice-oriented approach for gathering requirements for digital twins. Based on these particular requirements, we can derive a broader framework leading to the second research question:

RQ2: What are the fields of interest /categories for practitioners concerning the requirements for a digital twin?

From these synthesized requirements, modules and clusters bring further order to the field of knowledge. We expect multiple insights from interviews with industrial experts. However, as only few companies have implemented a digital twin, we must wonder whether the industry is ready to tackle the topic of digital twins:

RQ3: Is it worthwhile to have a closer look at the practical requirements of digital twins?

Knowledge about the requirements of a specific object permits further scientific artifacts, which may enable practitioners to implement more of these artifacts. Together with the question of whether the practitioners are ready for implementation, we pave the way for further practice-oriented research.

The paper is structured as follows. The following section provides the basics about digital twins, after which we outline the research approach. In section four, we discuss the results and outline the different requirements. Finally, we conclude the paper and provide the contributions, limitations, and further research possibilities.

2. Theoretical Background

Starting from physical twins that have been applied in production for decades [7], the concept of digital twins constantly evolves [2]. The original digital twin concept stems back to a lecture by Michael Grieves describing a digital twin as a three-part concept [8]. According to Grieves, a digital twin contains a physical product, a virtual product, and a bi-directional data flow between both products. This description was expanded by two additional elements, including sensor data and historical data sets at a later point [3]. [9] provides further relevant insights into digital twins, defining digital twins as an extension of digital models and distinguishing digital twins from digital shadows through the type of data linkage since digital twins possess a bi-directional data linkage to their physical counterparts. Furthermore, [10] describe the data handling methods of a digital twin and focus their work on the service domain of a digital twin. [2] and [11] provide more specific descriptions of digital twins. The former describes eight different dimensions of a digital twin: data link, purpose, conceptual elements, accuracy, interfaces, synchronization, data input, and time of creation [2]. The latter describes similar dimensions. Nevertheless, they add the dimensions physically, environment, fidelity, and state of the system [11].

[4] provide a combination of the above developments and define five archetypes of digital twins. The following definition forms the basic understanding of digital twins in this contribution:

> "The Digital Twin is a virtual construct that represents a physical counterpart, integrates several data inputs with the aim of data handling, data storing, and data processing, and provides an automatic, bi-directional data linkage between the virtual world and the physical one. Synchronization is crucial to the Digital Twin to display any changes in the state of the physical object. Additionally, a Digital

> Twin must comply with data governance rules and must provide interoperability with other systems" [4, p. 14].

Requirements for digital twins are seldomly examined. A comparison within the common databases shows that less than one percent of the literature on digital twins deals with requirements for digital twins [12]. The current works are either literature reviews (e.g., [13] or [14]) or focus on specific domains, mainly in the context of manufacturing (e.g., [15] or [16]). Commonly listed aspects are a bi-directional data linkage, specific interfaces (HMI and M2M), or real-time capabilities. Many requirements, however, may be directly synthesized from the given definitions. Therefore, we aim for a more practical approach and capture the requirements from real-world applications.

3. Research Approach

Requirements engineering is a methodology for defining requirements for a technical system or a technology such as the digital twin [17]. In this context, requirements are properties that technology needs to possess. These requirements can also be framework conditions under which the technology must perform [18]. According to [19], requirements engineering consists of the four steps identification, analysis, specification, and validation.

First, the requirements must be identified by analyzing existing systems or conducting interviews with selected experts. This leads to a plethora of unsorted requirements (step 1). We chose an empirical approach and conducted several interviews (see Table 1). This paper presents research in progress and, hence, we started with a small sample of six interviews, three with experts from logistics and three with experts from productional contexts. Another goal of this study is to evaluate whether more in-depth research with an extended interview series is appropriate. Hence, further interviews will follow.

Table 1: Sectors and Company Sizes of the Interviewees

#	Sector	Company Size
1	Logistics	>50,000 employees
2	Production	>80,000 employees
3	Production	>20,000 employees
4	Production	>7,000 employees
5	Logistics	>100,000 employees
6	Logistics	>50,000 employees

The interviews follow the approach of [20]. The starting point is the identification of the interview partners. All six interviewees are experts with many years of experience in their respective fields of work. The mixture of the interviewees provides a balanced picture of different domains, i.e., logistics and production. The interviews followed a semi-structured approach, as we did not prescribe many questions but aimed for a free conversation and just provided the interviewee with thematic blocks [21]. To identify the necessary expertise of the interviewee, we first asked introductory comprehension questions regarding digital twins. Nevertheless, following the advice of [21] and [22], we ensured combability between the individual discussions through minor guidance if an interviewee digressed from the core of the discussion, the requirements. All interviews were recorded and transcribed.

During the analysis's second step, we sorted the requirements into groups. These groups contain requirements, which show relations amongst them. One way to analyze is to differentiate between functional and non-functional requirements. According to [17], functional requirements define what is to be executed by a technology. On the other hand, non-functional requirements describe how the technology should function. Another task of the analysis is the elimination of identified redundancies, and lastly, the prioritization of the requirements (step 2). During the specification phase (step 3), we streamline the formatting of the requirements. We furthermore performed both steps simultaneously and listed all requirements chronologically. Then, we coded each requirement to gain comparability between the requirements. In addition, it allowed us to identify and synthesize duplicates. In total, the interviews provided 30 individual requirements. After the analysis and specification, eight distinctive requirements were left. Finally, there is a validation of the requirements by comparing the requirements with the stakeholders' expectations (step 4).

4. Requirements for Digital Twins

The eight distinctive requirements deal with different aspects of a digital twin, i.e., data handling, data policies, or digital twin services (DT Services). Table 2 provides an overview of the requirements.

Table 2: Derived Requirements

RQM	Requirement	Category
1	Synchronization between digital and physical parts must be reliable	Data Handling
2	Data sovereignty / Control over which data is exchanged and for how long	Data Policies
3	Data security, data protection, and data governance rules must be implied	Data Policies
4	Role allocation for digital twins through neutral standard access rules	Data Policies
5	A digital twin must possess real-time monitoring and data analytics	DT Services
6	A digital twin must possess simulation and prediction tools	DT Services
7	Data sharing capabilities and interfaces for data sharing must be present	Data Handling
8	A digital twin must be (at least) semi-automated	Data Handling

4.1 Requirement 1: Synchronization

Three out of six interviews demand specifically a reliable synchronization between a digital twin and its physical counterpart. In this context, all interviewees refer to real-time synchronization. Nevertheless, one interviewee (interview 4) preferred a semi-manual synchronization, allowing for a manual update of the digital twin and a fully automated solution. This is especially beneficial when dealing with use cases at an early stage of development, where the physical system is not yet equipped with devices for real-time communication, i.e., IoT-capable sensors.

The other two interviewees did not specify the level of automation of the synchronization. Nevertheless, both demanded a reliable synchronization, which draws a precise image of the real-time conditions:

“Now the components and requirements. […] It [the digital twin] should be easy to integrate and provide the transparency about the real real-time conditions.”	Interview 1

This corresponds with many descriptions from the literature, e.g., [3], [4], or [11]. Hence, real-time capable synchronization is a critical requirement for digital twins.

4.2 Requirement 2: Data Sovereignty

Five out of six interviewees consider data sovereignty and control over data a mandatory requirement for digital twins. Data sovereignty is crucial for digital business processes as it enables active monitoring of data usage and the restriction of unwanted use [23].

Data sovereignty is essential for data sharing via digital twins:

> "Interviewer: So, you are saying in this context there have to be certain requirements that you can restrict areas [of the digital twin], that you control access, or only release what you want to release and for how long you want to release it?
>
> Interviewee: Yes, absolutely, because otherwise, I possess a great model that I will never disclose to anyone because otherwise I am completely blank and lose my competitive advantage."

Interview 4

Interviewees agree that maintaining data sovereignty is fundamental for collaborative data sharing. However, as [24] state, digital twins form an essential basis for the collaborative use of technical data. Therefore, the consideration of data sovereignty is mandatory for digital twins. Based on the general importance of data security for industrial data sharing and the interviews, we see that the aspect of data sovereignty is of crucial significance for digital twins.

4.3 Requirement 3: Data Security, Data Protection, and Data Governance

Data governance and the related data security and protection mechanisms are the third requirement. Four interviewees agree with the importance of data governance for digital twins. However, the data governance may affect two aspects – the data and the digital twin:

> "There is data governance via the data sources. Where does the digital twin get its data? Then it's the digital twin itself. That is more or less a data representation."

Interview 2

Data governance rules for data are not new, as they are generally necessary and determine any data handling processes. Nevertheless, data governance rules for a digital twin are a novelty. It is not surprising that these governance rules are requested, as a digital twin, as any digital object, permits data manipulation and, hence, needs rules for data protection and safety.

4.4 Requirement 4: Role allocation

Related to the data sovereignty and data governance, but introduced as an independent aspect, is the role allocation:

> "A digital twin should be so modular that I have the option of making the digital twin or the functions of this twin accessible to certain user groups. So that it is a multiuser concept, with appropriate roles and permissions."

Interview 2

Half of the interviewees demand a role allocation. While this allocation is somewhat a standard procedure regarding data sharing and access tools, it is not typical concerning digital twins. The interesting aspect is that role allocation is demanded not only for the data inside the twin but also for the twin itself. Besides aspects like who may see data and contents of the digital twin, another essential aspect is the ownership and legal responsibility of the digital twin. This is particularly important for digital twins that accompany a specific product over its life cycle. The B2C sector, where the digital twin over all instances of the product line may stay with the manufacturer, does not seem to pose a considerable problem. Nevertheless, this circumstance may lead to significant disputes in the B2B sector. While company A sells the physical product to companies B and C, it wants to retain control of the digital twin. On the other hand, companies B and C are interested in possessing the digital twin data.

A rule catalog for these cases is mandatory. For example, the seller may offer instances of the digital twin, in which the buyer only accesses the data of his physical counterparts.

4.5 Requirements 5 & 6: Services of the Digital Twin

Recent works enhance the classical view of digital twins and demand that a digital twin provides certain services (e.g., see [25] or [26]). A service defines as a non-physical performance a company offers [27]. In this context, we define the services of a digital twin as possibilities to work with data within the twin.

Having a closer look at the literature about digital twins, simulation, prediction, monitoring, and analysis seem to be the most important services a digital twin offers [28,4]. The interviews back this expectation:

“I think one important point is a video analytics use case. For example, you could start using a "heat map" to make the topics and processes of employees transparent. You know that they are there today and are working well, and you have some characteristics or some functions that you can read from the scan points, but how do the processes run in detail, and what potential is there? These are the topics I would start with.”	Interview 1
“The digital twin is also a model, a shortened representation of reality and, in this case, a simulation that automatically shows the current status.”	Interview 4

The interviewees request these services as mandatory parts of a digital twin. Though, the interviewees distinguish between two specifications. On the one hand, the services monitoring and analysis represent one symbiosis, while on the other hand, the services simulation and prediction create the second aggregation. In both cases, a service that leverages the results of the preceding ones is merged. We follow the more realistic view and consolidate the different services into requirements five and six. A digital twin should contain monitoring and analysis as well as simulation and prediction services.

4.6 Requirement 7: Data Sharing and Interfaces

Besides the different services a digital twin provides, the interviewees requested data sharing capabilities and interfaces for data sharing.

“A shared digital twin is definitely a topic that is very, very interesting and very, very important for us that the digital twin runs across organizational boundaries. Because we are looking at the lifecycle here.”	Interview 5

As described by the interviewees, a digital twin needs data from different entities when monitoring the life cycle. Therefore, a digital twin must have appropriate interfaces that enable such distributed data acquisition. Furthermore, these interfaces are even more critical if a digital twin has to represent logistics processes. In this case, multiple data sources from different entities of the supply network must be synthesized.

4.7 Requirement 8: Semi-Automation

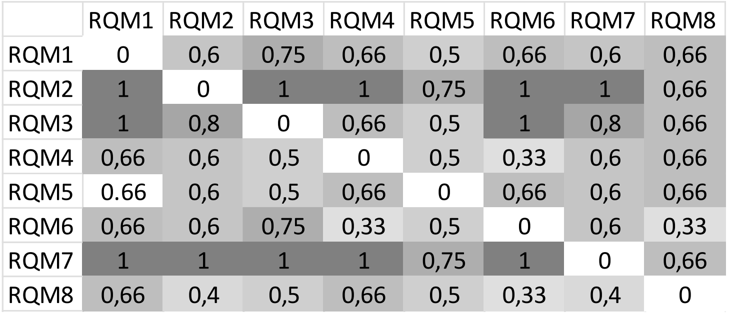
Lastly, the level of automation must be analyzed. While interviewee 2 demands a fully automated digital twin, this is impossible for every individual use case.

„For me, it is already a prerequisite that this process is automated. That information flows automatically into these systems in which or from which the digital twin obtains its information or even directly flows into the digital twin automatically. “	Interview 2

Many use cases demand a manual interference opportunity. Infrastructural use cases even need manual interference interfaces for regulatory reasons. Therefore, in this context, it is necessary to have the option of manual intervention. This leads to the necessity of a semi-automated digital twin. The requirement is closely related to the first one. However, as requirement one concentrates on the (automated) synchronization, here a broader focus must be applied, as all processes within a digital twin are included.

4.8 Dependencies amongst the Requirements

A closer look at the contents of the requirements shows similarities between several requirements. Self-evident are the requirements RQM 2 – 4 and RQM 6 / 7. The first group of requirements is related to each other. All three requirements deal with rules and concepts for data security. Summarized, we subsume these requirements under the category of data policies. Similarly, requirements six and seven both describe the services of a digital twin. Hence, these requirements are subsumed under digital twin (DT) services. As stated above, these two requirements are functional requirements, which underline the service character of these two. The remaining three requirements, RQM 1, 5, and 8, describe how data is handled within a digital twin. Thus, we merge these three requirements in the category of data handling. These categories are preliminary and can be expanded or combined as further requirements arise since categorization is highly dependent on individual use cases.



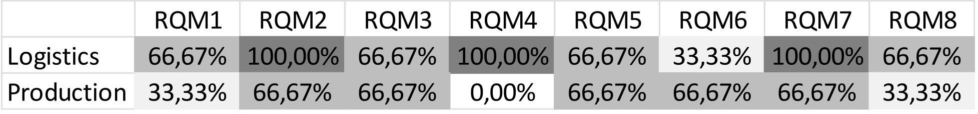
	RQM1	RQM2	RQM3	RQM4	RQM5	RQM6	RQM7	RQM8
RQM1	0	0,6	0,75	0,66	0,5	0,66	0,6	0,66
RQM2	1	0	1	1	0,75	1	1	0,66
RQM3	1	0,8	0	0,66	0,5	1	0,8	0,66
RQM4	0,66	0,6	0,5	0	0,5	0,33	0,6	0,66
RQM5	0.66	0,6	0,5	0,66	0	0,66	0,6	0,66
RQM6	0,66	0,6	0,75	0,33	0,5	0	0,6	0,33
RQM7	1	1	1	1	0,75	1	0	0,66
RQM8	0,66	0,4	0,5	0,66	0,5	0,33	0,4	0

Figure 1: Correlations of the Requirements

Figure 1 shows further dependencies between the requirements. The figure is read in columns. We compare how often a specific requirement is mentioned together with another one. For example, each interviewee who demanded RQM 1 also mentioned RQM 2, whereas only two-thirds mentioned RQM 4. Fields with 100% show requirements that should be implemented together. Very high values scores RQM 1 – synchronization – and RQM 6 – simulation and prediction. Digital twins, which are synchronized, therefore need data sovereignty, data governance, and data sharing capabilities. This seems obvious, though, data sovereignty and governance are often neglected in practice. For reliable results, these policies are whatsoever mandatory.

Similarly, a digital twin that performs predictions and simulations requires reliable data inputs, as a solid database is essential for simulation purposes [29]. The comparatively low values for RQM 5, especially in conjunction with RQM 1, are surprising. Like the solid database for simulation, we consider a reliable and, above all, up-to-date database essential for monitoring purposes, but only half of the interviewees agree with this.

Another aspect is the distribution of the requirements in comparison to the sectors. Very important for logistics seems to be data sovereignty, the allocation of roles to users, and data sharing capability. This is plausible because of the nature of logistics as a highly distributed discipline, with many interfaces and different participants within one ecosystem. Hence, digital twins for logistics should focus on data sharing and distributing data while keeping measures in place to protect the data and underlying metadata.



	RQM1	RQM2	RQM3	RQM4	RQM5	RQM6	RQM7	RQM8
Logistics	66,67%	100,00%	66,67%	100,00%	66,67%	33,33%	100,00%	66,67%
Production	33,33%	66,67%	66,67%	0,00%	66,67%	66,67%	66,67%	33,33%

Figure 2: Sectoral Distribution of the Requirements

A peak of interest for one requirement is not evident for productional contexts. Interestingly, only the role allocation does not seem to play any role for manufacturers. We see this justified through the relatively small sample of manufacturers. Nevertheless, the domain logistics shows that a more significant research study with more interviews might bring more precise insights.

5. Conclusion, Limitations, and Contributions

In this contribution, we aimed to identify requirements for digital twins in logistics and production. For this purpose, we conducted an interview series with experts from the industry. Their answers were coded and analyzed. Furthermore, eight requirements were derived (RQ1). Namely, these are synchronization between digital and physical parts, data sovereignty, data security, data governance, and data policies, role allocations for a user with access to the digital twin, services a digital twin provides, e.g., monitoring and simulation, interfaces for data sharing, and semi-automated processes.

Related to RQ2, these eight requirements may be grouped into different categories, i.e., data handling, data policies, and digital twin services. These categories provide further opportunities to analyze more requirements depending on each category. Lastly, this research should show whether further research is worthwhile (RQ3). The relatively small interview series provides very interesting insights. Hence, a broader interview series will provide deeper insights and may bring specific requirements for certain domains.

Our paper is subject to limitations. While we focused on the highest level of objectivity, subjective influences cannot be ruled out during the coding process. Furthermore, the study is relatively small. But to find out whether this research approach is likely to be successful, we accepted the small sample. The scientific contributions show ways for further research. Broader studies focusing on particular domains or subjects of the digital twin should be carried out. Furthermore, this research provides progress to the body of knowledge of digital twins concerning needs the research on digital twins must tackle. As managerial contributions, eight distinctive requirements for practical digital twins are provided. Practitioners can include them in the respective developments of their digital twins. Additionally, knowledge about the requirements might help during the design phase of a digital twin, which most companies are now. As requirements lay the foundation for multiple scientific artifacts, and besides the already mentioned broader study on the requirements itself, numerous opportunities for further research are possible. For example, design principles, reference models, or concrete implementations of digital twins are thinkable.

Acknowledgements

This research was conducted in the project “Silicon Economy Logistics Ecosystem” and is funded by the German Federal Ministry of Transport and Digital Infrastructure.

References

[1] Panetta, K., 2019. Gartner Top 10 Strategic Technology Trends for 2020. https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/. Accessed 20 November 2019.

[2] van der Valk, H., Haße, H., Möller, F., Arbter, M., Otto, B., 2020. A Taxonomy of Digital Twins, in: AMCIS 2020 Proceedings. AIS, Salt Lake City, USA.

[3] Grieves, M., Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems, in: Kahlen, F.-J., Flumerfelt, S., Alves, A. (Eds.), Transdisciplinary Perspectives on Complex Systems. New Findings and Approaches. Springer International Publishing, Cham, Switzerland, pp. 85–113.

[4] van der Valk, H., Haße, H., Möller, F., Otto, B., 2021. Archetypes of Digital Twins. Business & Information Systems Engineering.

[5] GE Power Digital Solutions, 2016. GE Digital Twin. Analytic Engine for the Digital Power Plant. http://www.ge.com/digital/sites/default/files/download_assets/Digital-Twin-for-the-digital-power-plant-.pdf. Accessed 31 January 2022.

[6] Siemens, 2018. MindSphere. Enabling the world's industries to drive their digital transformations. https://www.prolim.com/wp-content/uploads/2019/04/Siemens-MindSphere-Whitepaper.pdf. Accessed 31 January 2022.

[7] Rosen, R., Wichert, G. von, Lo, G., Bettenhausen, K.D., 2015. About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. IFAC-PapersOnLine 48 (3), 567–572.

[8] Grieves, M., 2014. Digital Twin: Manufacturing Excellence Through Virtual Factory Replication. Michael W. Grieves LLC.

[9] Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in Manufacturing: A Categorical Literature Review and Classification. IFAC-PapersOnLine 51 (11), 1016–1022.

[10] Tao, F., Zhang, H., Liu, A., Nee, A.Y.C., 2019. Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics 15 (4), 2405–2415.

[11] Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., 2020. Characterising the Digital Twin: A Systematic Literature Review. CIRP journal of manufacturing science and technology 29, 36–52.

[12] Scopus, 2022. Search "Requirements and Digital Twins". https://www.scopus.com/results/results.uri?sort=plf-f&src=s&sid=7998efb288aeea0386966f4847f516cb&sot=a&sdt=a&cluster=scolang%2c%22English%22%2ct&sl=38&s=TITLE%28requirements+AND+%22digital+twin%22%29&origin=searchadvanced&editSaveSearch=&txGid=1e031d5618f43c66a3daf880b188180f. Accessed 11 January 2022.

[13] Carvalho, R., Da Silva, A.R., 2021. Sustainability Requirements of Digital Twin-Based Systems: A Meta Systematic Literature Review. Applied Sciences 11 (12), 5519.

[14] Hajjem, E., Benderbal, H.H., Hamani, N., Dolgui, A., 2021. Digital Twin Framework for Reconfigurable Manufacturing Systems: Challenges and Requirements, in: Dolgui, A., Bernard, A., Lemoine, D., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems. IFIP WG 5.7 International Conference, APMS 2021, Nantes, France, September 5–9, 2021, Proceedings, Part II, vol. 631, 1st ed. 2021 ed. Springer International Publishing; Imprint Springer, Cham, pp. 553–562.

[15] Durão, L.F.C.S., Haag, S., Anderl, R., Schützer, K., Zancul, E., 2018. Digital Twin Requirements in the Context of Industry 4.0, in: Chiabert, P., Bouras, A., Noël, F., Ríos, J. (Eds.), Product Lifecycle Management to Support Industry 4.0, vol. 540. Springer International Publishing, Cham, pp. 204–214.

[16] Lehner, D., Pfeiffer, J., Tinsel, E.-F., Strljic, M.M., Sint, S., Vierhauser, M., Wortmann, A., Wimmer, M., 2021. Digital Twin Platforms: Requirements, Capabilities, and Future Prospects. IEEE Softw., 0.

[17] ISO/IEC/IEEE, 2018. Systems and software engineering — Life cycle processes — Requirements engineering. ISO/IEC/IEEE, Switzerland.

[18] Institute of Electrical and Electronics Engineers, 1990. IEEE Standard Glossary of Software Engineering Terminology.

[19] Pohl, K., 2010. Requirements engineering: Fundamentals, principles, and techniques. Springer, Berlin, Heidelberg, 813 pp.

[20] Sarker, S., Sarker, S., 2009. Exploring Agility in Distributed Information Systems Development Teams: An Interpretive Study in an Offshoring Context. Information Systems Research 20 (3), 440–461.

[21] Myers, M.D., Newman, M., 2007. The Qualitative Interview in IS Research: Examining the Craft. Information and Organization 17 (1), 2–26.

[22] Patton, M.Q., 2002. Qualitative Research & Evaluation Methods. SAGE Publications, Thousand Oaks, USA, London, UK, New Dehli, India.

[23] Jarke, M., Otto, B., Ram, S., 2019. Data Sovereignty and Data Space Ecosystems. Business & Information Systems Engineering 61 (5), 549–550.

[24] Haße, H., van der Valk, H., Weißenberg, N., Otto, B., 2020. Shared Digital Twins: Data Sovereignty in Logistics Networks, in: Proceedings of the HICL.

[25] Qi, Q., Tao, F., Zuo, Y., Zhao, D., 2018. Digital Twin Service towards Smart Manufacturing. Procedia CIRP 72, 237–242.

[26] Zhang, H., Ma, L., Sun, J., Lin, H., Thürer, M., 2019. Digital Twin in Services and Industrial Product Service Systems. Procedia CIRP 83, 57–60.

[27] Lies, J., 2018. Service. Gabler. https://wirtschaftslexikon.gabler.de/definition/service-42239/version-265590. Accessed 11 January 2022.

[28] Enders, M.R., Hoßbach, N., 2019. Dimensions of Digital Twin Applications - A Literature Review, in: Proceedings of the 25th Americas Conference on Information Systems, Cancun: Mexico, pp. 1–10.

[29] Law, A.M., 2015. Simulation Modeling and Analysis, 5th ed. McGraw-Hill Education, New York, NY.

Biography

Hendrik van der Valk (*1992) has been a researcher at the Chair for Industrial Information Management at the TU Dortmund University since 2019. He studied Mechanical Engineering, and his research focuses on Digital Twins in production and logistics. He conducts his research as a member of the Excellence Center for Logistics and IT and the Silicon Economy. He is currently a member of the Association for Information Systems AIS and the Institute for Operations Research and the Management Sciences INFORMS.

Jan-Luca Henning (*1996) studies Industrial Engineering at TU Dortmund University and has been working as a research assistant at the Chair for Industrial Information Management in Dortmund since 2019. He conducts his research as a member of the MINIS – Miniaturized Production and Logistics System project.

Stephanie Winkelmann (*1994) has been a researcher at the Chair for Industrial Information Management at the TU Dortmund University since 2022. She has graduated with a Bachelor's degree in Business Administration and Logistics and holds a Master of Science in Logistics from TU Dortmund University. In cooperation with the Fraunhofer Institute for Software and Systems Engineering and the Fraunhofer Institute for Material Flows and Logistics, she conducts her research as a member of the Silicon Economy.

Hendrik Haße (*1992) is researcher and did his PhD at the Fraunhofer Institute for Software and Systems Engineering. He holds a Master of Science in Industrial Engineering with a focus on Industrial Management from the TU Dortmund University. His research focus is on IoT architectures and the design of distributed Digital Twins. In cooperation with the Chair of Industrial Information Management at the TU Dortmund University he conducts his research on the conceptual structuring of Digital Twins.

3rd Conference on Production Systems and Logistics

Structuring The Digital Energy Platform Jungle: Development Of A Multi-Layer Taxonomy And Implications For Practice

Sebastian Duda[1], Lukas Bank[2], Can Kaymakci[3], Jana Köberlein[2], Simon Wenninger[1,4], Tobias Haubner[5], Alexander Sauer[3], Johannes Schilp[2]

[1]*Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Bayreuth, Germany*
[2]*Fraunhofer Institute for Casting, Composite and Processing Technology IGCV, Augsburg, Germany*
[3]*Fraunhofer Research Institution for Manufacturing Engineering and Automation IPA and Institute for Energy Efficiency in Production (EEP), Stuttgart, Germany*
[4]*FIM Research Center, University of Applied Sciences Augsburg, Germany*
[5]*Technical University Darmstadt, Germany*

Abstract

Rising and volatile energy prices are forcing production companies to optimize their consumption patterns and reduce carbon emissions to remain competitive. Demand-side management (DSM) or energy flexibility (EF) is a promising option for the active management of electricity demand. With DSM, energy procurement costs can be effectively reduced, for example, by reducing peak loads and taking advantage of volatile energy prices. In addition, renewable energies can be better integrated to reduce carbon emissions while stabilizing the power grid. Although the benefits of DSM for production companies are well known, implementation is not yet widespread. A key barrier is the high requirements of IT systems and the associated effort and complexity involved in setting them up. Companies often lack appropriate IT systems or have historically grown systems that do not allow continuous communication from the machine to the energy market. A variety of different platforms promise solutions to address these challenges. However, when selecting platforms, it is often unclear which aspects and functionalities of a platform are relevant for a company's specific application. To address this gap, we developed a multi-layer taxonomy of digital platforms for energy-related applications in the industry that includes a general, as well as a more specific data-centric and transaction-centric perspective. We develop, revise, and evaluate our taxonomy using insights from literature and analysis of 46 commercially available platforms or platforms developed through research projects. Based on our taxonomy, we derive implications for research and practice. Our results contribute to the descriptive knowledge of digital platforms in energy-related applications. Our taxonomy enables researchers and practitioners to classify such platforms and make informed decisions about their deployment.

Keywords

Digital Platform; Taxonomy; Demand-Side-Management; Energy Flexibility; IT Systems

1. Introduction

Adverse effects of human-induced climate change afford targeted and effective measures for achieving the climate goals set out in the international climate agreements [1]. The phase-out of coal and nuclear power generation was adopted as a key measure by the German government [2]. To meet electricity demand, the share of electricity generation from renewable energy sources is expected to increase to 80% of electricity consumption by 2050 [3]. This shift in the electricity generation portfolio will pose major challenges for power grids, and price volatility is expected to increase significantly [4]. The highly fluctuating, weather-dependent electricity generation from renewable energy sources such as wind and solar power, which is only adaptable to a limited extent, requires new solutions for a secure electricity supply. Besides the expansion

DOI: https://doi.org/10.15488/12192

of power grids, the increase of storage capacities, and the use of potentials of sector coupling, demand-side management (DSM) offers a competitive solution to address the expected challenges by increasing the energy flexibility of the demand side [5]. Considering that the industrial sector accounts for almost 44% of electricity consumption in Germany, it offers significant potential for balancing fluctuations in the power grid by adjusting electricity consumption to the electricity supply [6]. Typically, energy-intensive industrial companies can shut down, shift, or regulate their (production) processes and plants deviating from their regular use in order to adjust their electricity demand [7]. With DSM, companies can benefit from reduced energy procurement costs by responding to volatile electricity prices or lowering their grid charges by avoiding peak loads [6]. While companies benefit monetarily from DSM, they moreover support the integration of renewable energy by adjusting power consumption to minimize carbon emissions while stabilizing the power grid [8]. Although the benefits of DSM for production companies are well known, implementation is not yet widespread, and concerns exist [9]. A major obstacle to the implementation of DSM in production companies is the high requirements for IT systems and the associated expense and complexity of setting them up [10]. These high requirements result from information flows beyond company boundaries, the interaction of diverse optimization services, and automation through transparency and standardization of the entire process of energy flexibility marketing [6,11]. Companies often lack appropriate IT systems or have historically grown systems that do not allow them to meet the requirements [12,13]. In addition, continuous communication in an often-heterogeneous IT system landscape is not possible due to a lack of interfaces [6,13]. In the meantime, the market of available platforms has grown considerably, and there are a variety of different platforms that promise (partial) solutions to these challenges [14,10,13]. However, for companies that intend to implement DSM and need to select suitable platform solutions, it is often unclear which aspects and functions of a platform are relevant for them. The evaluation of available platforms is time-consuming, and tools and assistance such as a pre-classification of platforms and their characteristics do not exist. This study, therefore, aims to address this pertinent gap in research and practice.

Therefore, we develop a multi-layer taxonomy of digital platforms for DSM applications in industry that includes a general, as well as a more specific data-centric and transaction-centric perspective. For this purpose, we develop, revise, and evaluate our taxonomy following the iterative multi-step method of Berger et al. [15]. We use insights from literature and analysis of 46 commercially available platforms or platforms developed through research projects. Based on our taxonomy, we derive implications for research and practice contributing to the descriptive knowledge of digital platforms in DSM applications. Our taxonomy enables users to classify such platforms and make informed decisions about their deployment.

2. Background

2.1 Demand-Side Management

The aim of demand-side management (DSM) is the management of demand for grid-based services among consumers in industry, commerce, or private households [16]. DSM generally adjusts the energy demand without having to increase or decrease the energy supply and can therefore be a sufficient solution for the energy transition to decentralized and highly volatile electricity generation [17]. Especially the energy-intensive industries such as metal production, chemicals, or the paper industry offer high potential for industrial DSM since they are responsible for approximately two-thirds of the industrial electricity consumption in Germany [18]. Studies showed that the potential for DSM in energy-intensive processes, e.g., aluminum electrolysis, is remarkable [19]. In addition to reducing energy procurement costs, e.g., by minimizing peak loads or shifting electricity consumption to times with lower electricity prices, the use of flexibility can generate additional benefits and potential revenue streams, such as by offering ancillary services [6]. When operationalizing industrial DSM, the impact of flexibility measures and flexible processes on energy demand and the resulting influences on production systems and schedules must be considered in order to avoid a negative impact on logistical production goals [8]. Flexibility measures can be used at different production system levels and address different business areas and operating resources of a production company. The control of complex processes and flexibility measures therefore requires suitable IT systems that ensure transparency and enable the automation of DSM [6,11].

2.2 Energy Platforms

In recent years, digital platforms have emerged in many business areas to bring customers and providers together and offer innovative services [14]. The term platform is used very frequently, but its meaning is not clear and uniform [20]. IT platforms are already being used today for the digitization and optimization of production. Digital services such as predictive maintenance or the optimization of production planning are used [21]. However, most existing and commercially available IT platforms are tailored to the products and services offered by the respective provider. They tend to use proprietary rather than open interfaces and protocols, forming a closed ecosystem [22]. As a result, neither interactions with external systems nor interoperability with other platform providers are possible. In addition to traditional digital platforms for optimizing production processes regarding logistical targets, many providers offer software products in the area of energy management. These products optimize the energy flow in production processes [23]. Energy management platforms are mostly used for capturing, processing, and monitoring energy flows within a company. Also, decision support systems for energy procurement and optimization are established in the market. The offered services range from electricity market forecasts to solutions for production scheduling optimization considering electricity market prices [24]. To sum up, there are established platforms and IT systems with a strong focus on production (systems) and infrastructure, offering a wide range of possible solutions for companies [13]. Even though there is research on architectural features of industrial Internet of Things platforms [25], on taxonomies of products and platforms for energy feedback technologies [26], or energy-efficient resource management technique taxonomies in platforms as service clouds [27], there is to the best of the authors’ knowledge no taxonomy or structuring element in literature that focuses on digital platforms and IT systems for DSM applications in production companies. This study, therefore, aims to contribute to this research vacuum and support companies in practice.

3. Methodological Approach

To address the elaborated gap in research and practice, and to structure the complex and heterogeneous field of digital energy management platforms, the development of a taxonomy is the adequate method. Taxonomies, also referred to as frameworks or typologies, serve to classify objects according to their characteristics and help to better understand, analyze, and structure knowledge and objects [15]. A taxonomy contains various dimensions, which in turn consist of at least two (mutually exclusive) characteristics allowing objects to be classified according to their characteristics. Nickerson et al. [28] proposed a method for developing taxonomies in an iterative process that is now well established, frequently used, and further developed in business information systems research.

For our work, we apply the iterative multi-step method of Berger et al. [15], who extended the method of Nickerson et al. [28] to include validation after taxonomy development. We, therefore, determine the meta-characteristic as well as objective and subjective ending conditions, which serve as an orientation and basis for the taxonomy development, in a first step. In accordance with our research goal, we define the meta-characteristic as “key distinguishing features of energy platforms in terms of their structure and main functions.” We adopt the conditions given by Nickerson et al. [28] and Berger et al. [15] as objective ending conditions that must be checked for fulfillment after each iteration of taxonomy development: (1) each characteristic is unique within its dimension, (2) each dimension is unique and not repeated within the taxonomy, and (3) each object has been studied, (4) at least one object must be identified per characteristic and dimension, (5) the characteristics of a dimension are mutually exclusive, (6) no new dimensions or features have been added or changed in the last iteration. For the subjective final conditions, we also followed previous research and decided that the taxonomy must be assumed by all authors to be concise, robust, comprehensive, extendible, and explanatory [28]. In a second step, we develop our taxonomy in four iterative rounds of conceptual-to-empirical and empirical-to-conceptual approaches until all ending conditions are met. Nickerson et al. [28] allow for an iterative combination of conceptual-to-empirical and empirical-to-conceptual approaches during taxonomy creation. Within the conceptual-empirical approach, we elaborate dimensions and their characteristics incorporating findings extracted from a literature review on existing platform research. We thereby test the dimensions and characteristics identified by assigning digital energy platforms to them. For the empirical-conceptual approach, we build on existing commercially available platforms and those available within research projects. First, we group the platforms. Second, we inductively derive the taxonomies dimensions and characteristics. To do so, we conducted online research and identified 163 platforms as a first selection. These include IoT platforms and energy management platforms, as well as aggregators and platforms for trading energy and flexibility. Since it was not possible to include all 163 platforms during the taxonomy creation process, but all platform types still had to be considered equally. A stratified sample selection was performed according to Quatember [29]. In doing so, 46 different platforms were identified, which are listed in Table 1. In our work, we alternate between a conceptual-empirical and an empirical-conceptual approach until in the fourth iteration all ending conditions were met. While conducting the iterations, we found that mutual exclusion of the characteristics is not possible for some dimensions without having to forego relevant information. Therefore, we tried to keep this condition as far as possible and to omit it only for dimensions that absolutely require it. Similar findings have already been made in other works on taxonomy development [30,31].

Table 1: Considered platforms during taxonomy development

Platform type	Platforms
Trading platforms	Cordinet Project, Cornwall Local Energy Market, Electron Platform, ETPA, Flexible Power, FutureFlow, GoFlex, Nextra, Nodes Market, Piclo Flexibility Marketplace, wepower
IoT platforms	AWS IoT Core, Bosch IoT Suite, CELOS, Cloud der Dinge, Connected Factories, Connected Factories 2, Enterprise IoT Platform, FIWARE, Google IoT Core, IBM Watson, ITAC.MES.Suite, LITMUS, OpenIoTFog, Productive 4.0, PTC Thingworx, Siemens Mindsphere, tapio, Virtual FortKnox
Energy management platforms	Bosch Energy Platform, DEXMA Platform, EMPURON EVE, EnCoMOS, ennex OS, ITC Power Commerce EnMS, KMUPlus - Energy Intelligence, opti.node, PHI-Factory, SIMATIC Energy Suite, Smart Energy Hub
Aggregators	Balance Power, BayWa r.e. CLENS, Centrica Business Solutions, e2m, Entelios, Next-Kraftwerke

In a third step, after all, end conditions were met, we validated the developed taxonomy by conducting interviews with eleven collaborators from the Connected Factories, DEXMA, CELOS, Internet of Things, Litmus IoT, Nodes Market, PHI-Factory, Thingworx, SIMATIC Energy Suite, Smart Energy Hub, and tapio platforms and had them categorize their platforms into the taxonomy. No problems occurred during the validation, and all persons were able to fully classify their platforms into the taxonomies. Therefore, the taxonomy could finally be approved, and the creation process completed. We then discussed and derived implications for research and practice in the last step.

4. Results

Using this three-step approach, we identified 15 dimensions with their specific characteristics for digital energy platforms and subdivided them into general, data-centric, and transaction-centric dimensions.

4.1 General Dimension of Energy Platforms

The platforms examined are operated either by an independent company, by a consortium consisting of several companies and institutions, or by an aggregator (*platform operator*). An aggregator administrates a virtual power plant with energy flexibilities of several companies, thereby acts as a third party compared to the first two options. *Access* to the platforms is possible via a web app using any internet browser, a native app installed on hardware, or via specific programming interfaces, via which data can be imported, exported, and exchanged with other systems. Depending on the access design, stronger or weaker lock-in effects may occur. There are various options for the *operational concept*. In on-premise operation, the platform is operated by the customers on their own IT infrastructure. In this case, the control and management of all data lie entirely with the customers, and the platform can be operated and maintained independently of the providers once it is up and running [32]. Alternatively, the platform can be operated in the cloud, which includes private clouds and public clouds. The on-premises and cloud operating modes can also be combined into a mixed form, which is called hybrid. Customers can decide for themselves which data should be processed in the private sphere and which can be uploaded to the cloud [33]. If the platform offers free access, clients can register without restrictions. Often, however, certain criteria must be met. These are checked by the platform providers as part of a prequalification. Some platforms require the use of certain devices, e.g., specific hardware for data collection (*access requirements*). The *platform structure* also shows different characteristics. It is either fixed by the operator, modular without external interfaces, whereby customers can freely choose additional functions or extensions but are bound to the operator's offer, or modular with external interfaces, whereby the platform can be supplemented with external offers in addition to those of the operator.

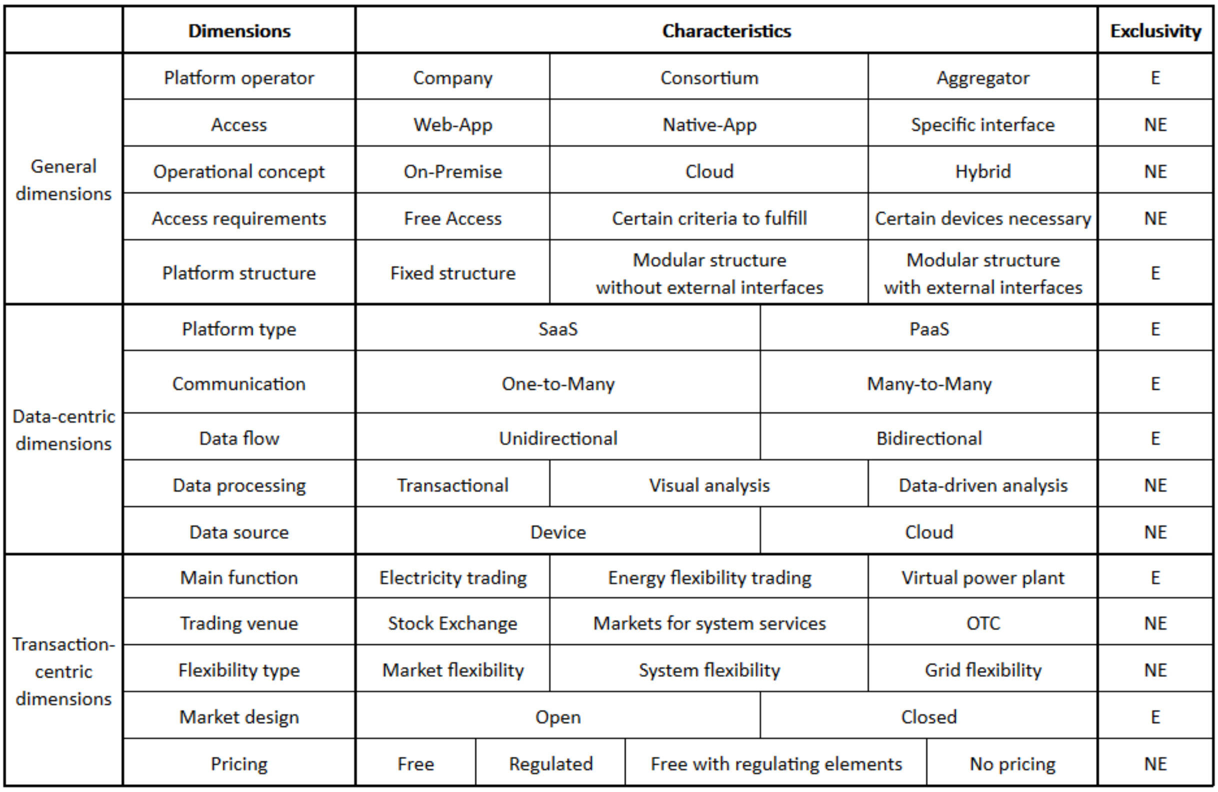
4.2 Data-Centric Dimensions of Energy Platforms

Platforms can correspond to two basic models (*platform type*) [34,32]. Software-as-a-Service (SaaS) refers to applications that can be used directly by the customers. In contrast, Platform-as-a-Service (PaaS) offers an environment in which applications can either be provided or developed. By definition, both models are offered via the cloud, whereby the required infrastructure is also provided by the service providers. However, since in practice, the on-premises operation is often also possible, and the operators themselves refer to their platforms as SaaS or PaaS, we use these terms in our characteristics. *Communication* via the platform may proceed either as one-to-many or many-to-many, depending on how the individual participants communicate with each other. Participants are the users or devices (systems, machines) connected to the platform. If several participants communicate exclusively with the platform, this is called one-to-many. In many-to-many communication, the platform not only communicates with several partners, but these also communicate with each other [35]. The *data flow* is also characterized by two different features. Either the data only flows in one direction, for example, from the devices to the platform (unidirectional), or the data flow takes place in more than one direction (bidirectional), for example, from the devices to the platform and vice versa. This also includes data flow between individual devices. *Data processing* can be either transactional or analytical. If the data is processed transactionally, data from transactions or interactions is used to trigger certain processes. Analytical data processing is divided into two different characteristics, namely visual analysis and data-driven analysis [31]. The former contains descriptive analyses, which mainly aim at preparing the data. The feature data-driven analysis, on the other hand, contains more in-depth forms of data analysis, where the data is used for further calculations, such as in machine learning applications [36,37]. The processed data is gained from different sources (*data sources*). Many platforms offer the possibility to connect devices to it to process their data. However, data can also be obtained from the cloud, where data either comes from external sources (e.g., energy price forecasts) or from the company itself (e.g., from enterprise-resource-planning systems).

4.3 Transaction-centric Dimensions of Energy Platforms

The dimension *main function* addresses the fact that platforms can pursue a different central objective. In the case of transaction-centric energy platforms, we found that this is either electricity trading, where producers and consumers can sell or buy electricity (via over-the-counter (OTC) trading or exchange access), energy flexibility trading, which enables generators or consumers to market energy flexibility directly to grid operators or, virtual power plants, that bundle (decentralized) electricity consumers and electricity producers, market their generated electricity and offer system services. Stock exchanges (e.g., EPEX SPOT), markets for system services, or OTC trading can serve as a *trading venue* on these platforms. Different *flexibility types* are traded on these trading venues. Market flexibility can be used to respond to market signals such as volatile prices on short-term markets to reduce energy costs. System flexibility serves to maintain electricity grid stability and is therefore used by the transmission system operator, whereas grid flexibility is intended to avoid critical situations in the local electricity grid [38]. If the *market design* of a platform is closed, the platform users are bound to a specific buyer, an aggregator, a fixed trading venue, or to the platform itself, and parallel use of several solutions is not possible. In an open market design, there is no lock-in, and customers can use additional solutions. According to [39], energy platforms use three different mechanisms to set prices (*pricing*) (for marketing energy flexibility). In free pricing, prices are formed without restrictions. In the case of free pricing with regulatory elements, there are restrictions imposed by the platform operators, for example, by imposing surcharges on freely formed prices or by setting price caps. In regulated pricing, prices are formed according to set procedures, or there are fixed prices. However, some platforms do not allow for their own price formation but only pass on prices from certain trading venues.

4.4 A Multi-Layer Taxonomy for Digital Energy Platforms



	Dimensions	Characteristics				Exclusivity
General dimensions	Platform operator	Company	Consortium	Aggregator		E
	Access	Web-App	Native-App	Specific interface		NE
	Operational concept	On-Premise	Cloud	Hybrid		NE
	Access requirements	Free Access	Certain criteria to fulfill	Certain devices necessary		NE
	Platform structure	Fixed structure	Modular structure without external interfaces	Modular structure with external interfaces		E
Data-centric dimensions	Platform type	SaaS	PaaS			E
	Communication	One-to-Many	Many-to-Many			E
	Data flow	Unidirectional	Bidirectional			E
	Data processing	Transactional	Visual analysis	Data-driven analysis		NE
	Data source	Device	Cloud			NE
Transaction-centric dimensions	Main function	Electricity trading	Energy flexibility trading	Virtual power plant		E
	Trading venue	Stock Exchange	Markets for system services	OTC		NE
	Flexibility type	Market flexibility	System flexibility	Grid flexibility		NE
	Market design	Open	Closed			E
	Pricing	Free	Regulated	Free with regulating elements	No pricing	NE

Figure 1: Combined use of the developed taxonomies with general, data-centric, and transaction-centric dimensions

Energy platforms are often characterized by either a strong data-centric or a transaction-centric focus. The main function of transaction-centric platforms is to serve as a marketplace. Data-centric platforms focus on processing data [40]. Since transaction-centric and data-centric platforms differ significantly in some respects, we developed two different taxonomies for digital energy platforms. Thus, all platforms can be classified exactly and according to the requirements of a taxonomy. The taxonomy consists of the general combined with the data- or transaction-centered dimensions, regarding the analyzed platform. Among the energy platforms, there are some that, based on their main functions, can be classified as transaction- or data-centered platforms, but at the same time also have individual functions of the other category such as the platform solution illustrated in [6]. To classify holistic platforms by using our taxonomies the transaction- and data centric taxonomies can be united to a combined taxonomy as shown in Figure 1~~figure 1~~. The individual characteristics are indicated for each dimension in the corresponding row. The item exclusivity also indicates whether the characteristics are mutually exclusive (E) or non-exclusive (NE).

5. Implications and Conclusion

This paper addressed the lack of methodology to characterize digital energy platforms. Following Berger et al. [15], we developed a taxonomy for digital energy platforms. Our result, the developed and validated taxonomy, has several implications for practice and research. First, the taxonomy serves as a structuring element and allows companies to determine the status quo of existing platforms and IT systems. Therefore, the taxonomy can be used for "auditing purposes" to examine which functions and features are already covered by existing IT systems. Second, in addition to the status quo, the taxonomy can also sharpen a target image for functionalities and characteristics of IT systems and platforms and elicit requirements. Third, using a fit-gap analysis [41] and our taxonomy, companies can easily compare existing digital energy platforms and select the platform that best meets their needs. The selected platform can then be implemented to accelerate the adoption of DSM. Fourth, the results and the comparison of the platforms used during development show that most of the platforms focus on the data-centric or transaction-centric dimensions and cover these functionalities. Consequently, it can be concluded that for the best possible digitalization and automation of flexibility marketing, a mix of different IT systems and platforms is evident in most cases. Considering existing IT systems and platforms, attention must therefore be paid to integration interfaces during implementation to ensure communication without media discontinuity. Only in this way can systems be operated in an interoperable manner. Fifth, researchers can classify their work using our taxonomy and clearly distinguish it from existing studies. Consequently, the taxonomy can serve to structure research and identify future research fields. Sixth, to sum up, our taxonomy provides researchers and practitioners with an easy-to-use methodology to classify digital energy platforms and make informed decisions about which platform best fits the needs of the business. This enables researchers to strengthen the focus of their research and helps companies leverage the potentials of DSM.

Naturally, our work has some limitations as any research endeavors but likewise gives prospects for further research. First, the taxonomy was developed with platforms known today and consequently with their characteristics. Against the background of the rapid development of platforms of various types, the developed taxonomy may only be used for a limited time, and further development will be necessary for the future. In this way, not yet considered functions and features can be considered to provide a valid taxonomy. Second, the taxonomy was built upon the analysis of a subset of all available energy platforms. In total we identified 163 digital energy platforms. Using the method of Quatember [29] we built a subset of 46 platforms to create our taxonomy. This reduction of the set might affect the completness of the taxonomy. Third, in addition to the pure classification of platforms, the taxonomy might be used in further research to derive platform archetypes allowing to group platforms with similar characteristics. Third, the taxonomy dimensions are at a relatively high level and do not, for example, take into account details regarding the exact interfaces of platforms. Consequently, our taxonomy is suitable for an initial platform selection. Future studies might build on our work and develop a more detailed taxonomy. A focus on technical details might thereby enhance readability and simplify its use. Despite these limitations, we hope to provide a viable solution to structure digital energy platforms for production companies and support researchers and practice.

Acknowledgments

The authors gratefully acknowledge the financial support of the *Kopernikus*-Projekt *SynErgie* by the *Federal Ministry of Education and Research* (*BMBF*) and the project supervision by the project management organization *Projektträger Jülich* (*PtJ*).

References

[1] Rusche, S., Rockstuhl, S., Wenninger, S., 2021. Quantifizierung unternehmerischer Nachhaltigkeit in der Fertigungsindustrie: Entwicklung eines zielorientierten Nachhaltigkeitsindex. Z Energiewirtsch.

[2] Bundesministerium für Wirtschaft und Energie, 2019. Kommission "Wachstum, Strukturwandel und Beschäftigung" Abschlussbericht, Berlin.

[3] Umweltbundesamt, 2019. Erneuerbare Energien in Zahlen. https://www.umweltbundesamt.de/themen/klima-energie/erneuerbare-energien/erneuerbare-energien-in-zahlen#textpart-1. Accessed 5 September 2019.

[4] Bachmann, A., Bank, L., Bark, C., Bauer, D., Blöchl, B., Brugger, M., Buhl, H.U., Dietz, B., Donnelly, J., Friedl, T., Halbrügge, 2021. Energieflexibel in die Zukunft – Wie Fabriken zum Gelingen der Energiewende beitragen können.

[5] Schilp, J., Bank, L., Köberlein, J., 2021. Executive Summary: Konzept der Energiesynchronisationsplattform – Diskussionspapiere.

[6] Bauer, D., Hieronymus, A., Kaymakci, C., Köberlein, J., Schimmelpfennig, J., Wenninger, S., Zeiser, R., 2020. Wie IT die Energieflexibilitätsvermarktung von Industrieunternehmen ermöglicht und die Energiewende unterstützt. HMD.

[7] Lindner, M., Wenninger, S., Fridgen, G., Weigold, M., 2022. Aggregating Energy Flexibility for Demand-Side Management in Manufacturing Companies – A Two-Step Method, in: Behrens, B.-A., Brosius, A., Drossel, W.-G., Hintze, W., Ihlenfeldt, S., Nyhuis, P. (Eds.), Production at the Leading Edge of Technology. Springer International Publishing, Cham, pp. 631–638.

[8] Bank, L., Wenninger, S., Köberlein, J., Lindner, M., Kaymakci, C., Weigold, M., Sauer, A., Schilp, J., 2021. Integrating Energy Flexibility in Production Planning and Control - An Energy Flexibility Data Model-Based Approach.

[9] Cardoso, C.A., Torriti, J., Lorincz, M., 2020. Making demand side response happen: A review of barriers in commercial and public organisations. Energy Research & Social Science 64, 101443.

[10] Good, N., Ellis, K.A., Mancarella, P., 2017. Review and classification of barriers and enablers of demand response in the smart grid. Renewable and Sustainable Energy Reviews 72, 57–72.

[11] Wenninger, S., Kaymakci, C., Wiethe, C., Römmelt, J., Baur, L., Häckel, B., Sauer, A., 2022. How Sustainable is Machine Learning in Energy Applications? – The Sustainable Machine Learning Balance Sheet. 17th International Conference on Wirtschaftsinformatik.

[12] Kaymakci, C., Wenninger, S., Sauer, A., 2021. A Holistic Framework for AI Systems in Industrial Applications, in: Ahlemann, F., Schütte, R., Stieglitz, S. (Eds.), Innovation Through Information Systems, vol. 47. Springer International Publishing, Cham, pp. 78–93.

[13] Roesch, M., Bauer, D., Haupt, L., Keller, R., Bauernhansl, T., Fridgen, G., Reinhart, G., Sauer, A., 2019. Harnessing the Full Potential of Industrial Demand-Side Flexibility: An End-to-End Approach Connecting Machines with Markets through Service-Oriented IT Platforms. Applied Sciences 9 (18), 3796.

[14] Donnelly, J., John, A., Mirlach, J., Osberghaus, K., Rother, S., Schmidt, C., Voucko-Glockner, H., Wenninger, S., 2021. Enabling The Smart Factory – A Digital Platform Concept For Standardized Data Integration.

[15] Berger, S., Denner, M.-S., Roeglinger, M., 2018. The Nature of Digital Technologies - Development of a Multi-Layer Taxonomy. 26th European Conference on Information Systems (ECIS).

[16] Murthy Balijepalli, V.S.K., Pradhan, V., Khaparde, S.A., Shereef, R.M., 2011 - 2011. Review of demand response under smart grid paradigm, in: ISGT2011-India. 2011 IEEE PES Innovative Smart Grid Technologies - India (ISGT India), Kollam, Kerala, India. 01.12.2011 - 03.12.2011. IEEE, pp. 236–243.

[17] Müller, T., Möst, D., 2018. Demand Response Potential: Available when Needed? Energy Policy 115, 181–198.

[18] Ausfelder, F. (Ed.), 2018. Flexibilitätsoptionen in der Grundstoffindustrie: Methodik, Potenziale, Hemmnisse : Bericht des AP V.6 "Flexibilitätsoptionen und Perspektiven in der Grundstoffindustrie" im Kopernikus-Projekt "SynErgie - synchronisierte und erngieadaptive Produktionstechnik zur flexiblen Ausrichtung von Industrieprozessen auf eine fluktuierende Energieversorgung", 1. Auflage ed. DECHEMA Gesellschaft für Chemische Technik und Biotechnologie e.V, Frankfurt am Main, 296 pp.

[19] Sauer, A., Abele, E., Buhl, H.U. (Eds.), 2019. Energieflexibilität in der deutschen Industrie: Ergebnisse aus dem Kopernikus-Projekt - Synchronisierte und energieadaptive Produktionstechnik zur flexiblen Ausrichtung von Industrieprozessen auf eine fluktuierende Energieversorgung (SynErgie). Fraunhofer Verlag, Stuttgart, 746 Seiten.

[20] Reuver, M. de, Sørensen, C., Basole, R.C., 2018. The Digital Platform: A Research Agenda. Journal of Information Technology 33 (2), 124–135.

[21] Zhong, R.Y., Xu, X., Klotz, E., Newman, S.T., 2017. Intelligent Manufacturing in the Context of Industry 4.0: A Review. Engineering 3 (5), 616–630.

[22] Wajid, U., Bhullar, G., 2019. Towards Interoperability Across Digital Manufacturing Platforms, in: Popplewell, K., Thoben, K.-D., Knothe, T., Poler, R. (Eds.), Enterprise Interoperability VIII, vol. 9. Springer International Publishing, Cham, pp. 81–91.

[23] Lee, D., Cheng, C.-C., 2016. Energy savings by energy management systems: A review. Renewable and Sustainable Energy Reviews 56, 760–777.

[24] Siano, P., 2014. Demand response and smart grids—A survey. Renewable and Sustainable Energy Reviews 30 (3), 461–478.

[25] Arnold, L., Jöhnk, J., Vogt, F., Urbach, N., 2021. A Taxonomy of Industrial IoT Platforms' Architectural Features, in: Ahlemann, F., Schütte, R., Stieglitz, S. (Eds.), Innovation Through Information Systems, vol. 48. Springer International Publishing, Cham, pp. 404–421.

[26] Karlin, B., Ford, R., Squiers, C., 2014. Energy feedback technology: a review and taxonomy of products and platforms. Energy Efficiency 7 (3), 377–399.

[27] Piraghaj, S.F., Dastjerdi, A.V., Calheiros, R.N., Buyya, R., 2017. A Survey and Taxonomy of Energy Efficient Resource Management Techniques in Platform as a Service Cloud, in: Sugumaran, V., Chen, J.“., Zhang, Y., Gottschalk, R. (Eds.), Handbook of Research on End-to-End Cloud Computing Architecture Design. IGI Global, pp. 410–454.

[28] Nickerson, R.C., Varshney, U., Muntermann, J., 2013. A method for taxonomy development and its application in information systems. European Journal of Information Systems 22 (3), 336–359.

[29] Quatember, A., 2015. Datenqualität in Stichprobenerhebungen. Springer Berlin Heidelberg, Berlin, Heidelberg.

[30] Jöhnk, J., Roeglinger, M., Thimmel, M., Urbach, N. How to Implement Agile IT Setups: A Taxonomy of Design Options, in: , Proceedings of the 25th European Conference on Information Systems (ECIS).

[31] Roeglinger, M., Püschel, L., Schlott, H. What's in a Smart Thing? Development of a Multi-layer Taxonomy, in: , Proceedings of the 37th International Conference on Information.

[32] Reinheimer, S. (Ed.), 2018. Cloud Computing: Die Infrastruktur der Digitalisierung. Springer Vieweg, Wiesbaden, 216 pp.

[33] Vikas, S., Gurudatt, K., Vishnu, M., Prashant, K., 2013. Private Vs Public Cloud. International Journal of Computer Science & Communication Networks (2), 79–83.

[34] Jaekel, M., 2020. Disruption durch digitale Plattform-Ökosysteme: Eine kompakte Einführung. Springer Vieweg, Wiesbaden, Heidelberg, 112 pp.

[35] Porter, M.E., Heppelmann, J.E., 2015. How Smart, Connected Products Are Transforming Companies. Harvard Business Review, 96–114.

[36] Kaymakci, C., Wenninger, S., Pelger, P., Sauer, A., 2022. A Systematic Selection Process of Machine Learning Cloud Services for Manufacturing SMEs. Computers 11 (1), 14.

[37] Kaymakci, C., Wenninger, S., Sauer, A., 2021. Energy Anomaly Detection in Industrial Applications with Long Short-term Memory-based Autoencoders. Procedia CIRP 104, 182–187.

[38] Conexio GmbH, 2019. Zukünftige Stromnetze: 30.-31. Januar 2019, Novotel Berlin, Am Tiergarten, Berlin : Tagungsunterlagen. Conexio GmbH, Pforzheim, Deutschland, 510 pp.

[39] Radecke, J., Hefele, J., Hirth, L., 2019. Markets for Local Flexibility in Distribution Networks: Working Paper. ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg.

[40] Engelhardt, S., Wangler, L., Wischmann, S. Eigenschaften und Erfolgsfaktoren digitaler Plattformen. iit-Institut für Innovation und Technik in der, Berlin. https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/autonomik-studie-digitale-plattformen.pdf?__blob=publicationFile&v=6. Accessed 18 January 2022.

[41] Pajk, D., Kovačič, A., 2013. Fit Gap Analysis – The Role of Business Process Reference Models. Economic and Business Review 15 (4).

Biography

Sebastian Duda is working at the FIM Research Center of the University of Bayreuth and at the Project Group Business & Information Systems Engineering of the Fraunhofer FIT and focuses on research in the field of convergence of digital technologies.

Lukas Bank is working at Fraunhofer Institute for Casting, Composite and Processing Technology IGCV and focuses on Digital Twins in energy flexible factories. Applications from the field of optimization and simulation are also related to this.

Can Kaymakci has been working as a research associate in the Industrial Energy Systems Department at Fraunhofer IPA since January 2020. In the group Energy Flexible Production and Energy Data Analysis, he is involved in research and industrial projects using AI and Machine Learning for energy data.

Jana Köberlein is working at Fraunhofer Institute for Casting, Composite and Processing Technology IGCV. In the group sustainable production systems, her research focuses on the analysis and planning of flexible energy use in production as well as industrial energy supply concepts.

Simon Wenninger is working at the FIM Research Center of the University of Applied Sciences Augsburg and at the Project Group Business & Information Systems Engineering of the Fraunhofer FIT. He focuses on research in the fields of data analytics in an industrial and energy context.

Tobias Haubner is studying industrial engineering with a technical specialization in mechanical engineering at the Technical University of Darmstadt. Before, he was studying in the Management & Technology program at the Technical University of Munich.

Prof. Dr.-Ing. Alexander Sauer is the Executive Director of the Institute for Energy Efficiency in Production (EEP) at the University of Stuttgart, as well as Director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA.

Prof. Dr.-Ing. Johannes Schilp is Head of the Department of Processing Technology at Fraunhofer Institute for Casting, Composite and Processing Technology IGCV and holds the Chair of Production Informatics at the University of Augsburg.

3rd Conference on Production Systems and Logistics

A Framework For Structuring Resilience And Its Application To Procurement

Maria Linnartz[1], Günther Schuh[1], Volker Stich[1]
[1]*Institute for Industrial Management (FIR) at RWTH Aachen, Aachen, Germany*

Abstract

Companies operate in an increasingly volatile environment where different developments like shorter product lifecycles, the demand for customized products and globalization increase the complexity and interconnectivity in supply chains. Current events like Brexit, the COVID-19 pandemic or the blockade of the Suez canal have caused major disruptions in supply chains. This demonstrates that many companies are insufficiently prepared for disruptions. As disruptions in supply chains are expected to occur even more frequently in the future, the need for sufficient preparation increases. Increasing resilience provides one way of dealing with disruptions. Resilience can be understood as the ability of a system to cope with disruptions and to ensure the competitiveness of a company. In particular, it enables the preparation for unexpected disruptions. The level of resilience is thereby significantly influenced by actions initiated prior to a disruption. Although companies recognize the need to increase their resilience, it is not systematically implemented. One major challenge is the multidimensionality and complexity of the resilience construct. To systematically design resilience an understanding of the components of resilience is required. However, a common understanding of constituent parts of resilience is currently lacking. This paper, therefore, proposes a general framework for structuring resilience by decomposing the multidimensional concept into its individual components. The framework contributes to an understanding of the interrelationships between the individual components and identifies resilience principles as target directions for the design of resilience. It thus sets the basis for a qualitative assessment of resilience and enables the analysis of resilience-building measures in terms of their impact on resilience. Moreover, an approach for applying the framework to different contexts is presented and then used to detail the framework for the context of procurement.

Keywords

Resilience; Framework; Disruptions; Resilience Principles; Procurement

1. Introduction

Companies and their supply chains have frequently experienced different kinds of disruptions. Especially the COVID-19 pandemic as a recent example of a global crisis has caused major challenges in supply chains. According to a study by GYA ET AL. 80 % of the companies surveyed were affected and the pandemic resulted for example in shortages of critical materials, delayed deliveries and longer lead times, and difficulties in planning and adjusting production capacity to meet fluctuating demand occurred [1]. Disruptions are especially critical when they affect the procurement side [2]. A study by BVL demonstrated that disruptions impacts on the supply side have been considered worse [3]. Despite the frequency of disruptions, several developments lead to increasing complexity in supply chains making it more difficult to create transparency [4]. In this volatile environment, companies are not sufficiently prepared for disruptions. One way of dealing with disruptions is increasing resilience. Resilience enables a system to cope with even

DOI: https://doi.org/10.15488/12194

unexpected disruptions and to ensure competitiveness. [5,6] Building resilience is largely dependent on measures taken before a disruption occurs [7]. However, currently many companies do not systematically increase their resilience even though the importance of resilience is widely acknowledged [8]. As resilience is a multidimensional and complex construct, implementing resilience requires an understanding of its constituent parts. However, a common understanding is currently lacking. [9,10] It is thus necessary to decompose the multidimensional concept into its individual parts and understand the relationships between them. This paper aims at structuring resilience by developing a framework that strengthens the understanding of resilience. It, therefore, supports the analysis and design of resilience. Additionally, the proposed framework is applied to the context of procurement. The framework serves as a basis for a qualitative resilience assessment which is a prerequisite for systematically designing resilience. The remainder of this paper is organized as follows: Section 2 reviews the literature regarding resilience and existing frameworks for analyzing it. Section 3 presents the developed framework and section 4 summarizes the application of the framework to the context of procurement. Section 5 gives an outlook on the use of the framework.

2. State of the art

In this section, the term resilience in the context of supply chains and procurement is defined first, before existing frameworks for analyzing resilience are summarized and the research need is presented.

2.1 Definition of resilience

As stated before, resilience is a multidimensional concept that is used in various contexts like ecology, psychology, supply chain and economy. Currently, there is no common definition of resilience in the fields of supply chains and organisational resilience. [6] ANNARELLI AND NONINO distinguish between static and dynamic organisational resilience. While static resilience focuses on preventive actions to decrease the impact of disruptions, dynamic resilience aims at reactions and fast recovery [11]. According to ALI ET AL. supply chain resilience definitions differ in terms of the phases considered, the strategies covered and the abilities addressed. Phases of resilience include the periods before, during and after a disruption. Strategies that are covered in the definitions are proactive, reactive and concurrent. To distinguish between the abilities that are considered within the definitions ALI ET AL. identify the abilities to anticipate, to adapt, to respond, to recover and to learn. Based on these constructs, supply chain resilience definitions can be divided into narrower definitions that include only individual aspects of the phases, strategies and abilities and wider definitions that imply all phases and strategies. [9] This work chooses a comprehensive view and understands **resilience** as the ability of a company to prepare for potential disruptions, react and adapt to disruptions as well as the ability to return to the original state or achieve a better state after the disruption. The aim is to minimize the impact of disruptions through preventive measures and return to the original state as quickly and cost-effectively as possible. A **disruption** in this context is a temporary impact on the performance caused by the occurrence of a disruption event [12].

2.2 Resilience frameworks

The frameworks reviewed in this section can be divided into different categories: frameworks related to the resilience triangle, the disruption profile or both and frameworks related to resilience capabilities. The **resilience triangle** by BRUNEAU ET AL. is a quantitative measurement tool for the seismic resilience of communities. For measuring resilience the authors examine the performance development over time after a disruption and identify the time needed for recovery, the severity of the disruption measured as the drop in performance and the area between the original and the actual performance as resilience dimensions. [13] MELNYK ET AL. analyze the transient response of a system and build on a profile similar to the resilience triangle. The authors identify the time between the occurrence of the disruptive event and the disruptive effect, the time at which recovery sets in, the comparison between the original performance level and the performance level after recovery, and the area between the original and the actual performance level as relevant resilience dimensions. [14] SHEFFI AND RICE propose a **disruption profile** that describes the development of the performance over time and identify eight phases that characterize this profile [15]. MUNOZ AND DUNBAR apply the resilience triangle and refer to the disruption profile to observe the response behavior of actors in a supply chain. In addition to the original dimensions of the resilience triangle, they identify the profile length and the weighted sum as two dimensions that explicitly take into account the curve progression. [16] BEVILACQUA ET AL. combine the resilience triangle and the disruption profile to categorize the eight phases of the disruption profile for the supply chain context [17]. Frameworks that refer to **capabilities** qualitatively analyze resilience. PETTIT ET AL. distinguish between vulnerabilities and capabilities when analyzing resilience. While vulnerabilities decrease resilience, capabilities enhance resilience. [18] ALI ET AL. develop a framework that contains five capabilities and 13 corresponding elements. Additionally, the elements are detailed in practices that support building resilience. [9] DUCHEK focuses on organizational resilience and proposes a framework containing anticipation, coping and adaption capabilities. [10] The framework developed by GIANCOTTI AND MAURO includes five resilience phases and corresponding capabilities [19].

2.3 Research needs and requirements for the resilience structuring framework

The various definitions and different frameworks underline that there is no common understanding of resilience. Starting from the definitions, different authors set various focal points and thus include different aspects in their definitions. Additionally, existing frameworks take different perspectives. In the cases where authors identify specific components of resilience in their frameworks, these components differ across the analyzed frameworks. The frameworks considered are not sufficiently detailed to analyze the contribution of concrete resilience increasing measures. Moreover, the frameworks are not applied to the context of procurement. Overall, the analysis of existing approaches illustrates the lack of understanding of the building blocks of resilience. Thus, the remainder of this paper focuses on developing a framework that captures the different aspects of resilience. The framework needs to identify specific components which constitute resilience and especially take into account the different phases of resilience. Additionally, the components need to be structured to incorporate their interrelationships. For each component, the target direction for increasing resilience must be identified. This is especially important as the framework sets the basis for analyzing resilience increasing measures regarding their contribution to resilience. The developed framework should be generic and applicable to different contexts for example to procurement.

3. Development of a framework for structuring resilience

The proposed framework builds on the existing frameworks and contains different components that characterize resilience. First, a structure for the framework is proposed based on existing definitions of resilience that contains three main component groups. Then, the individual components for each group are identified by analyzing the existing approaches. Lastly, resilience principles which are understood as target directions for the systematic design of resilience are derived based on the identified components.

3.1 Structure of the framework

The analysis of existing definitions and frameworks demonstrates the dynamic aspect of resilience. When analyzing resilience different phases are of importance. In general, resilience influences the time before, during and after a disruption. [9] The first category of components, therefore, are **time-related components**. Additionally, resilience is defined through an intensity aspect [20]. Resilience focuses on the impact of disruptions on performance. Thus, the second category comprises **performance-related components**. These dimensions are also the basis of the above-described framework groups “resilience triangle” and “disruption profile”. As the disruption profile characterizes the time-related aspects and considers the performance, it will serve as a basis for identifying and framing the resilience components. Following the resilience triangle, the area between the original performance and the actual performance after a disruption can be used to characterize the resilience of the system considered. The smaller the area, the more resilient the system is. This area is not only influenced by the time and performance amounts but also by the progression of performance over time which is represented by the curve trajectory. The third category thus describes **curve-related components**. The structure of the framework is presented in Figure 1.

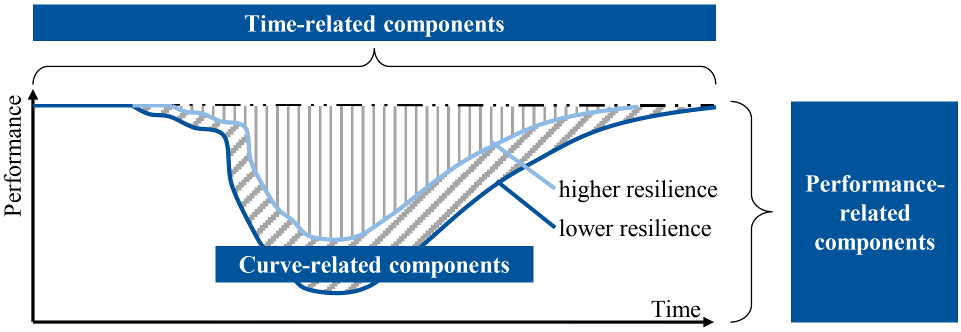


Figure 1: Framework structure

Time-related aspects specify the x-axis, whereby performance-related aspects define the y-axis and curve-related aspects characterize the curve progression. The analysis of the area in terms of opportunities for its reduction serves as a basis for identifying **resilience principles**. They indicate how the resilience components need to be modified to achieve higher resilience. Thereby, two overarching goals of resilience can be distinguished. On one hand, passive resilience aims at increasing the robustness of a system. A high robustness results in the least possible impact in the case of a disruption. On the other hand, active resilience aims at fast adaption and recovery when a disruption impact occurs. [21,22] The specific components and the corresponding resilience principles are described in the following.

3.2 Time-related components and resilience principles

For describing resilience the time before, during and after a disruption are important. Within this framework, time-related components refer to periods in the disruption profile which are defined by a start and endpoint. Before characterizing the time-related components it is, therefore, necessary to identify the relevant points in time. HEIL has characterized disruptions in terms of their time aspect and has identified several points in time that can be distinguished when analyzing a disruption [23]. A major point in time is the **occurrence of the disruption event (t_0)**. For the considered system the **start of the disruption impact (t_1)**, which can be observed through a decrease in performance, as well as the **end of the disruption (t_4)** are further relevant points. [23] Depending on the system state the occurrence of the disruption event and the start of the disruption can be separated points in time. The end of the disruption is characterized by reaching the original performance level in case of a full recovery or through reaching a new state of equilibrium in case of a partial recovery. For analyzing resilience components the point in time where it is known that the disruption event has occurred (**knowledge of the disruption event (t_2)**) is important as well. The last important point in time is the **lowest point of the performance curve (t_3)** as it marks the beginning of the recovery.

The first resilience component is defined as **buffer time**. A main aspect of resilience is concerned with minimizing the impact a disruption has on a system and ideally staying on the original performance level [21,16]. This component relates to the latent disruption phase as it is described by HEIL. It starts when the disruption event occurs and ends with the beginning of the disruption impact. [23] The existing system structure prevents a negative influence of the disruption. These structures are already present before the disruption occurs. Thus, the buffer time directly relates to robustness. The buffer time reflects the absorptive capacity of a system which ensures that the performance does not decrease even after a disruption event has happened. To increase the resilience this time should be as long as possible. The corresponding resilience principle is thus the **buffer time extension**.

Resilience is significantly characterized through the reaction in case of a disruption. In this context, the period that passes before recovery starts is an important feature of resilience. [6,14] The second component is thus called **response time**. In this framework, the response time includes the time that is needed to choose and implement specific actions. As a distinction to the buffer time, this component is linked to a concrete interference in the system. The response time begins with the knowledge of the disruption event. It ends when the measures take effect and the performance increases (the lowest point of the performance curve). The point in time when knowledge of the occurrence of the disruption event exists can be detailed when it is set in relation to other points in time. When looking at one system, the earliest point where the disruption can be known is the occurrence of the disruption event. However, it depends on the system characteristics if the disruption event is discovered right away. It is known at the latest when the disruption impact takes place. Depending on when the disruption event is discovered, the response time and the buffer time can overlap. The length of the response time is influenced by different latencies that can occur between an event and the effectiveness of countermeasure. These latencies contain the time that is needed to discover the event, analyze it, decide on the measures taken and the time for the measures to be effective [24]. Moreover, the response time depends on the available resources within the system. [6]

The response time corresponds to two resilience principles. First, the response time helps to increase the resilience when it is as short as possible. The shorter the response, the faster the lowest performance level is reached and recovery starts. Thus, the first resilience principle is **response time reduction**. Additionally, the start of the response time influences resilience. As the response time includes the time needed to choose measures and the time needed for these measures to become effective, an early start of the response results in earlier effective measures. The second resilience principle corresponding to the response time is therefore **response start shortening**.

The last time-related component refers to the recovery which occurs after measures take effect. This is also one of the two central aspects of the resilience triangle [13]. **Recovery time** is therefore proposed as the third component. The recovery time starts when the performance curve starts increasing. The end of the recovery time depends on whether a full or a partial recovery occurs. As described above, in the case of a full recovery the end is characterized by the fact that the original performance level is reached. For partial recovery, the end is determined by reaching a new equilibrium state of performance. The recovery time is characterized by an increase in performance caused by effective measures that have been taken. Like the response time, the recovery time leads to an increase in resilience when it is as low as possible. The corresponding resilience principle is thus called **recovery time reduction**.

3.3 Performance-related components and resilience principles

As described above, performance-related components refer to changes in performance that occur due to a disruption. They demonstrate the intensity of the disruption impact. Within the proposed framework, performance-related components are characterized through the difference between the original and the actual performance level during the disruption at a certain point in time. To identify relevant components, the extreme values in the curve are analyzed.

One important aspect is the maximum impact a disruption has on a system. This characterizes the severity of a disruption [13,16]. Following the resilience triangle and the disruption profile, the **maximum performance reduction** is defined as the first component. A high resilience is reached if the maximum performance reduction is as low as possible. This target dimension can be directly derived from the resilience triangle. Thus, this component relates to the resilience principle **damping of the maximum performance reduction**.

Additionally, the long-term performance level after the recovery has taken place is important [14]. As described above, either a full recovery or a partial recovery is possible. In the case of a partial recovery, the performance level after the recovery time has reached a new equilibrium state that is below the original level. Thus, the second component is defined as the **long-term performance reduction**. This aspect is also considered within the disruption profile [15]. This component characterizes the recovery capacity of a system. Similar to the above-described resilience principle, the target direction for the long-term performance reduction is its decrease. The optimum is reached if no long-term performance reduction exists. The corresponding resilience principle is therefore called the **damping of the long-term performance reduction**.

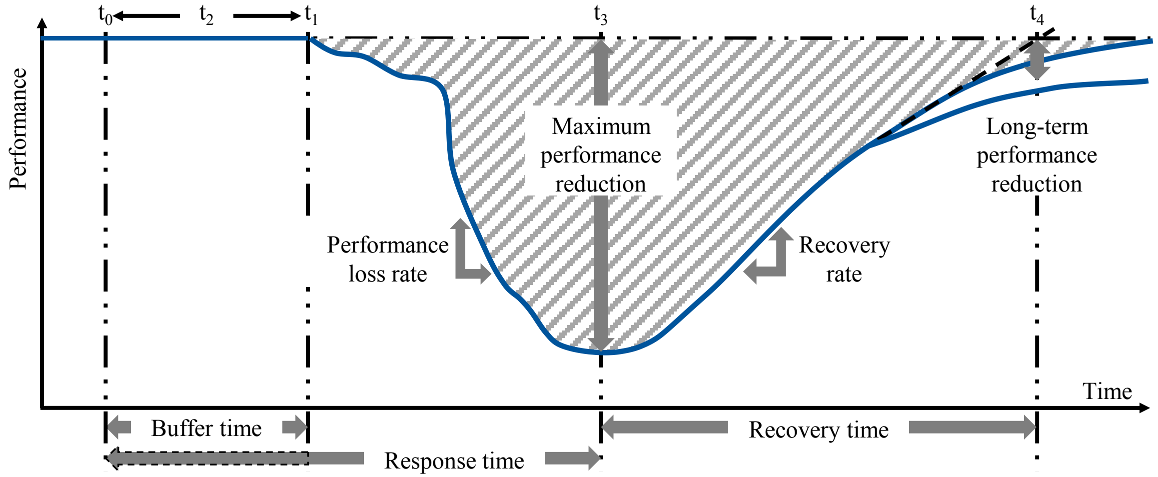
3.4 Curve-related components and resilience principles

The curve progression influences the area between the original and the actual performance level both during the reaction and the recovery time. To take into account different kinds of curve progressions, Munoz and Dunbar define the weighted sum and the length of the profile as two factors for measuring resilience. [16] Both factors are significantly influenced by the gradient of the curve. This corresponds with the approach of Cimellaro et al. who describe different recovery functions like linear, exponential and trigonometric to distinguish between different kinds of recoveries [25]. For the curve-related components of the proposed framework, the focus thus lies on the gradient of the curve.

The first curve-related component is defined as the **performance loss rate**. It refers to the gradient of the curve during the response time when the performance decreases. A high performance loss rate results in a high performance decrease in a short amount of time. This component is both influenced by the characteristics of the disruption as well as by the response behavior of the system. A low performance loss rate results in a slow and possible controlled reduction of performance. High resilience is thus reached if the performance loss rate is as low as possible. This results in the resilience principle **performance loss rate reduction**.

Despite the time required for recovery and the performance level reached after the recovery, the recovery of a system is characterized through the rate at which recovery takes place. The second component is thus defined as the **recovery rate**. The recovery rate is influenced by the effectiveness of the measures taken. In contrast to the performance loss rate, the recovery rate should be high to ensure a high resilience. This component thus corresponds to the resilience principle **recovery rate increase**.

Figure 2 shows the developed framework with the identified resilience components.



$\mathbf{t_0}$: occurrence of disruption event, $\mathbf{t_1}$: start of the disruption impact, $\mathbf{t_2}$: knowledge of disruption event, $\mathbf{t_3}$: lowest point of the performance curve, $\mathbf{t_4}$: end of the disruption

Figure 2: Resilience framework with components

Resilience can be characterized through seven components in the dimensions of time, impact and curve progression and the eight corresponding resilience principles. By using the disruption profile as the underlying structure, the framework takes into account the relationship between the identified factors. Additionally, the consideration of the area in combination with the described resilience principles details the multidimensional construct of resilience and points out various targets for systematically configuring resilience.

4. Application of the framework to procurement

The developed framework is generic and thus applicable for different contexts. In the following, an approach for applying the framework to different contexts is presented and illustrated for the area of procurement

The disruption profile is significantly influenced through the disruption, its impact and the different points in time. These aspects depend on the system under consideration. Therefore, the proposed framework applies to a specific object where a disruption impact can occur. For each object, a separate disruption profile has to be considered. However, often there are interactions between different objects. This is enhanced as disruptions are often complex and characterized by cause-effect chains that result in multiple level disruptions. Thus, a disruption impact happening at one object can cause a disruption impact at a linked object. In the framework, this is represented by looking at several disruption profiles that are linked by transition times. The link can concern both a physical material flow and an information flow. A link that concerns the material flow can lead to cascading disruption impacts if the system does not recover in time. A link that affects the information flow between two objects can influence the start of the response time. As described above, when looking at one object the response time starts between the occurrence of the disruption event and the start of the disruption impact. When taking into account several objects, the response time can already start when knowledge about a disruption occurs in a linked object if it is known that the disruption in the earlier stage will also affect the considered object at a later stage. For this to be effective, the information needs to be passed on between the objects.

To apply the framework to a specific context, a three-step approach is proposed: First, the relevant **performance values** must be characterized. The performance values specify the parameters whose changes in the disruption course are considered. Then, the **objects under consideration** need to be identified by analyzing the existing resources and possible disruption impacts. Lastly, when looking at several objects the **transition times** need to be specified.

For the context of procurement, the relevant **performance values** are identified in the following. The overall goal of resilience is to minimize the negative aspect a disruption has on the performance. Thus, the focus lies on the disruption impacts rather than on the disruption event or the disruption source. This is especially suitable as a large number of disruption sources leads to a limited number of impacts [26]. For the application of the framework to the context of procurement potential disruption impacts as well as overall goals of procurement are analyzed. Potential disruption impacts in procurement that are mentioned in the literature are a lack of material [26], business interruptions [27], deviations from the expected quality, planned quantity, planned delivery date or planned price [28], supply at the wrong time or supply of the wrong products [29] and differences between the desired and the actually supplied quantity [30]. Looking at these impacts it becomes apparent that when considering the physical material flow the impacts relate to missing material. Thus, *material availability* is chosen as the performance value. Material availability is the supply of the right material, at the right time, in the right quality, in the right quantity, at the right place [31]. This performance value corresponds with the overall goal of procurement. Namely, to ensure the long-term supply security of the enterprise for the production of goods [32].

The **objects** for which a disruption profile needs to be considered and their **interdependencies** are derived from the structure and resources in procurement. The procurement function links manufacturers and suppliers. Typical actors within a procurement chain include suppliers, manufacturers, logistical service providers and distributors [33]. The application of the framework in the context of procurement focuses on manufacturers. As the *goods receipt* is the direct interface to the suppliers or the logistical service providers, it serves as the central object that is considered. Material availability is influenced if the material is not supplied at the planned delivery date. The *production* is the internal demand source of the goods receipt. It is thus defined as the second object. A deviation in material availability occurs when input material for the production is not available at the planned date. To display the information linkage between the goods receipt and the external supply market, a third object is considered. This object comprises suppliers and logistical service providers as they both execute deliveries to the manufacturer. This object is called *external input actor*. Material availability at this object influences the ability to deliver. Goods receipt and external input actors are linked by procurement logistics while intralogistics links goods receipt and production. The transition times are the time needed for the transport respectively the time needed for the provision of material. The resulting resilience model is summarized in Figure 3.

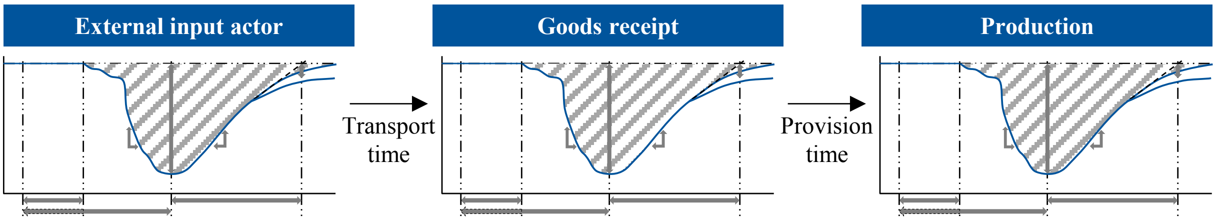


Figure 3: Resulting resilience model for the context of procurement

Each of these objects can experience deviations in material availability as a disruption impact. Additionally, an object can act as a disruption source for a subsequent object. For example, if a supplier is unable to deliver due to a production machine failure, the performance in the disruption profile of the external input actor decreases. Depending on the time needed for recovery, this can cause a performance decrease in the disruption profile of the goods receipt, if the material is not supplied at the planned delivery. Contingent on what resilience measures are in place at the goods receipt, a disruption impact can occur at the production.

5. Summary and outlook

Resilience is a multidimensional construct where understanding of its constituent parts is currently lacking. Such an understanding is the necessary foundation for analyzing and designing resilience. Therefore, a framework for structuring resilience has been developed. The framework builds on the disruption profile and the resilience triangle and contains seven resilience components. Additionally, it proposes eight resilience principles as target dimensions for building resilience. The generic framework can be applied to different contexts by defining the relevant performance values and the objects that need to be considered. This has been demonstrated for procurement. The framework thus sets the basis for systematically analyzing resilience improving measures regarding their specific contribution to resilience. Based on these results resilience can be configurated systematically. The paper contributes to an understanding of resilience, especially in the context of procurement. Further research is needed to identify concrete measures that increase resilience in procurement and analyze their contribution to the identified resilience principles.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Gya, R., Lago, C., Becker, M., Junghanns, J., Petit, J.-P., Perea, L., Schneider-Maul, R., Dahlmeier, S.C., Kumar, V., Penka, A., Buvat, J., KVJ, S., Nambiar, R., Puttur, R.K., 2020. Fast Forward: Rethinking the supply

chain resilience for a post-COVID-19 world, 44 pp. https://www.capgemini.com/de-de/wp-content/uploads/sites/5/2020/11/Fast-forward_Report.pdf. Accessed 31 December 2020.

[2] Pereira, C.R., Christopher, M., Lago da Silva, A., 2014. Achieving supply chain resilience: the role of procurement. Supply Chain Management: An International Journal 19 (5/6), 626–642.

[3] BVL.digital (Hrsg.), 2020. Digitale Erfolgsfaktoren für resiliente Wertschöpfungsketten: Studienbericht, Bremen, 49 pp. https://www.bvl.de/misc/filePush.php?id=53373&name=BVL_Digital_Presseinfo_Studie_Supply_Chain_Risk_management_201021.pdf. Accessed 17 March 2021.

[4] Huth, M., Romeike, F., 2016. Grundlagen des Risikomanagements in der Logistik, in: Huth, M., Romeike, F. (Eds.), Risikomanagement in der Logistik. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 13–47.

[5] Biedermann, L., 2018. Supply Chain Resilienz: Konzeptioneller Bezugsrahmen und Identifikation zukünftiger Erfolgsfaktoren. Springer Gabler, Wiesbaden.

[6] Kamalahmadi, M., Parast, M.M., 2016. A review of the literature on the principles of enterprise and supply chain resilience: Major findings and directions for future research. International Journal of Production Economics 171, 116–133.

[7] Kiebler, L., Ebel, D., Klink, P., Sardesai, S., 2020. Risikomanagement disruptiver Ereignisse in Supply Chains. Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund. https://www.iml.fraunhofer.de/content/dam/iml/de/documents/OE%20220/Risikomanagement_disruptiver_Ereignisse_in_Supply_Chains.pdf. Accessed 18 March 2021.

[8] van Hoek, R., 2020. Research opportunities for a more resilient post-COVID-19 supply chain – closing the gap between research findings and industry practice. International Journal of Operations & Production Management 40 (4), 341–355.

[9] Ali, A., Mahfouz, A., Arisha, A., 2017. Analysing supply chain resilience: integrating the constructs in a concept mapping framework via a systematic literature review. Supply Chain Management: An International Journal 22 (1), 16–39.

[10] Duchek, S., 2020. Organizational resilience: a capability-based conceptualization. Bus Res 13 (1), 215–246.

[11] Annarelli, A., Nonino, F., 2016. Strategic and operational management of organizational resilience: Current state of research and future directions. Omega 62, 1–18.

[12] Fischäder, H., 2007. Störungsmanagement in netzwerkförmigen Produktionssystemen, 1. Aufl. ed. DUV Deutscher Universitäts-Verlag, Wiesbaden.

[13] Bruneau, M., Chang, S.E., Eguchi, R.T., Lee, G.C., O'Rourke, T.D., Reinhorn, A.M., Shinozuka, M., Tierney, K., Wallace, W.A., Winterfeldt, D. von, 2003. A Framework to Quantitatively Assess and Enhance the Seismic Resilience of Communities. Earthquake Spectra 19 (4), 733–752.

[14] Melnyk, S.A., Zobel, C.W., Macdonald, J.R., Griffis, S.E., 2014. Making sense of transient responses in simulation studies. International Journal of Production Research 52 (3), 617–632.

[15] Sheffi, Y., Rice, J.B., 2005. A Supply Chain View of the Resilient Enterprise. MITSloan Management Review 47 (1), 41–48.

[16] Munoz, A., Dunbar, M., 2015. On the quantification of operational supply chain resilience. International Journal of Production Research 53 (22), 6736–6751.

[17] Bevilacqua, M., Ciarapica, F.E., Marcucci, G., 2017. Supply Chain Resilience Triangle: The Study and Development of a Framework. International Scholarly and Scientific Research & Innovation 11 (8), 1992–1999.

[18] Pettit, T.J., Fiksel, J., Croxton, K.L., 2010. Ensuring Supply Chain Resilience: Development of a Conceptual Framework. Journal of Business Logistics 31 (1), 1–21.

[19] Giancotti, M., Mauro, M., 2020. Building and improving the resilience of enterprises in a time of crisis: from a systematic scoping review to a new conceptual framework. 307-339 Pages / Economia Aziendale Online -, Vol 11, No 3 (2020): ISSUE 3/2020.

[20] Heinicke, M., 2017. Resilienzorientierte Beurteilung von Produktionsstrukturen. Dissertation, Magdeburg, 152 pp.

[21] Burnard, K.J., Bhamra, R., 2019. Challenges for organisational resilience. Continuity & Resilience Review 1 (1), 17–25.

[22] Wieland, A., Wallenburg, M.C., 2012. Dealing with supply chain risks. International Journal of Physical Distribution & Logistics Management 42 (10), 887–905.

[23] Heil, M., 1995. Entstörung betrieblicher Abläufe. Deutscher Universitätsverlag, Wiesbaden.

[24] Hackathorn, R., 2002. The BI Watch: Minimizing Action Distance. DM Review (12), 22–23.

[25] Cimellaro, G.P., Reinhorn, A.M., Bruneau, M., 2010. Seismic resilience of a hospital system. Structure and Infrastructure Engineering 6 (1-2), 127–144.

[26] Carvalho Remigio, H., 2012. Modelling resilience in supply chain. Dissertation, Lissabon, 196 pp.

[27] Sanchis, R., Poler, R., 2014. Enterprise resilience assessment: a categorisation framework of disruptions. Dirección y Organización 54, 45–53.

[28] Cube, J.P. von, Härtel, L., Schmitt, R., 2016. Model-based Decision Support in Supply Chains – Requirements for Monetary Supply Risk Quantification. Procedia CIRP 57 (7), 171–176.

[29] Datta, P., 2017. Supply network resilience: a systematic literature review and future research. The International Journal of Logistics Management 28 (4), 1387–1424.

[30] Jain, N., Girotra, K., Netessine, S., 2021. Recovering Global Supply Chains from Sourcing Interruptions: The Role of Sourcing Strategy. Manufacturing & Service Operations Management 46 (3), 75.

[31] Gottmann, J., 2019. Produktionscontrolling. Springer Fachmedien Wiesbaden, Wiesbaden.

[32] Piontek, J., 2016. Beschaffungscontrolling. De Gruyter Oldenbourg.

[33] Beckmann, H., 2019. Beschaffungslogistik, in: Furmans, K., Kilger, C. (Eds.), Betrieb von Logistiksystemen. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 19–105.

Biography

Maria Linnartz, M. Sc. (*1993) is a research associate at the Institute for Industrial Management (FIR) at the RWTH Aachen in the research group Supply Chain Management. She studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. In her research, she focuses on resilience and risk management in supply chains.

Prof. Dr.-Ing. Dipl.-Wirt. Ing Günther Schuh (*1958) holds the Chair of Production Systems at the Machine Tool Laboratory (WZL), is a member of the Board of Directors at the Fraunhofer Institute for Production Technology (IPT), Director of the Institute for Industrial Management (FIR) at the RWTH Aachen University and head of the Production Technology Cluster. He is the founder of the Schuh & Co. group of companies based in Aachen, St. Gallen and Atlanta.

Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

Industrial Human Activity Prediction and Detection Using Sequential Memory Networks

Tadele Belay Tuli[1], Valay Mukesh Patel[1], Martin Manns[1]
[1]*PROTECH - Institut für Produktionstechnik, Universität Siegen, Germany*

Abstract

Prediction of human activity and detection of subsequent actions is crucial for improving the interaction between humans and robots during collaborative operations. Deep-learning techniques are being applied to recognize human activities, including industrial applications. However, the lack of sufficient dataset in the industrial domain and complexities of some industrial activities such as screw driving, assembling small parts, and others affect the model development and testing of human activities. The InHard dataset (Industrial Human Activity Recognition Dataset) was recently published to facilitate industrial human activity recognition for better human-robot collaboration, which still lacks extended evaluation. We propose an activity recognition method using a combined convolutional neural network (CNN) and long short-term memory (LSTM) techniques to evaluate the *InHard* dataset and compare it with a new dataset captured in a lab environment. This method improves the success rate of activity recognition by processing temporal and spatial information. Accordingly, the accuracy of the dataset is tested using labeled lists of activities from IMU and video data. A model is trained and tested for nine low-level activity classes with approximately 400 samples per class. The test result shows 88% accuracy for IMU-based skeleton data, 77% for RGB spatial video, and 63% for RGB video-based skeleton. The result has been verified using a previously published region-based activity recognition. The proposed approach can be extended to push the cognition capability of robots in human-centric workplaces.

Keywords

Human-robot collaboration; human activity recognition; deep-learning; InHard dataset

1. Introduction

Today human-robot collaboration (HRC) is becoming an essential part of the industry for achieving better quality products in less time. In this regard, robots' cognition capabilities are expected to be enriched with the prediction and detection of human activities [1]. Such an approach that helps recognize human actions and activities may enable robots to complement human motions and activities in shared smart workplaces [2,3]. Lists of human activities in HRC may require semantical descriptions and definitions, which can be described either at a higher level with generalized actions or at a lower level with detailed descriptions. Generalized actions refer to, e.g., reach, pick, put, turn, and assemble. In contrast, low-level actions describe details such as reaching with the left hand to object A, taking a type B screwdriver, and tightening the screw. Prediction and detection of human action and sequence of activities, either higher or lower, is still basic research requiring further investigation.

DOI: https://doi.org/10.15488/12144

The classical approach in HRC task planning most often consists of pre-programmed logic that remains fixed for the given cycle operations. In a typical assembly, the robot executes the desired task in collaboration with the human operator in a programmed sequence. It hardly responds to changes in human performance which is one of the main barriers to HRC. Most recently, industries have been focusing on improving human and robot interaction in a shared workplace to complete a task efficiently and safely with flexibility in production processes. However, such a level of HRC is challenging as it involves many unpredictable events and actions, which are difficult for robots to understand and act accordingly. In this aspect, the robot must possess a cognitive capability and ability to understand various actions performed by the human operator to predict the subsequent possible action in order to carry out the task.

This work aims to present the activity recognition method using combined convolutional neural network (CNN) and long short-term memory (LSTM) techniques to evaluate the *InHard* dataset compared to a newly captured dataset in a lab environment.

2. Related Works

The current research that focuses on human activity recognition for HRC can be discussed from the aspects of types of human activities in industrial environments, methods employed for recognizing human actions, and generated datasets for human activity recognition (HAR).

2.1 Human activity types for HRC.

In today's industry 4.0 era, many researchers have brought humans and robots closer to the industrial environment. In [4], a detailed overview of HRC in industry 4.0 regarding various levels of humans and robots working together is given by solving safety issues using a particular collaborative robot (cobots). Furthermore, how the HRC can improve the efficiency of the industrial process by eliminating the uncomfortable, repetitive work of human operators is discussed. Similarly, [5] has conducted a survey to test the HRC process by measuring trust between humans and robots in an open workspace executing pick and place tasks. Work has to detail discussion about various safety factors and trust factors of HRC that can affect the productivity and efficiency of the process. Further, [6] has presented an HRC from a technical point of view. Various methods for human intention estimation through machine learning algorithms, robot action planning, and human-robot joint action planning are discussed and compared. A more detailed scenario of the industry is presented by [7], and it gives detailed information about the various industrial activities such as assembly activities, tool handling activities, and non-deterministic activities which are non-reparative, such as repairing activity or inspection, and also demonstrate that fusion of inertial measuring unit (IMU) sensors and video-based tracking system can be used to capture these activities with high precision. Similarly, [8] has also presented a work which includes modeling of industrial activities using a fusion of various motion capture sensors. It provides the detailed information of small industrial activities such as handling of nuts and bolts for assembly of the product and also to model various hand gestures movements to control the robot action. A visual sensor-based approach e.g., red-green-blue-depth (RGB+D) cameras have been also employed to capture various human activities in the industrial environments. Some of these activities include entering, leaving a work cell (movement), pointing to an object, waving (gesture), picking, and moving parts (object handling), applying pressure, reach to an object [8–10].

2.2 Methods for human activity recognition

Methods that have been employed for HAR can be considered into two big categories. The first is a statistical model, and the second is a deep learning-based model for activity prediction and detection.

Statistical models are known for their data-intensive requirement in order to generate high-quality motions [8]. Common approaches utilizing statistical models include Gaussian Mixture Models (GMM) or space partitioning (e.g., using k-means or principal component analysis (PCA) based linear mapping). GMM is a probabilistic model that maximizes expectation by fitting mixtures of Gaussian models to samples in high dimensional spaces. Deep learning-based models have been implemented for modeling human activities [11–13] based on video or skeleton. Open pose for two-dimensional pose estimation presented in [14] has employed a multi-stage CNN for extracting spatial features for human action recognition. Deep convolutional generative adversarial networks (GAN) have also been used to classify human activities even for fewer training datasets [15]. Further investigations for human motion generation or synthesis for enabling human interaction with smart machine systems that may involve higher-level human intention prediction and detection and lower-level details of actions have been shown in [16,17].

2.3 Human activity recognition datasets.

Dataset for HAR that publicly available includes HMDB51 [18], UCF50 [19], NTU RGB+D dataset [20], MSR-Action3D [21], and InHARD [10]. HMDB51 was introduced by [18], consisting of approximately seven thousand realistic video clips from sources such as movies and web series. The dataset consists of 51 classes of general day-to-day life activities such as jumping, laughing, kissing, and others, with 100 samples in each category. Another dataset in a similar category is UCF50 [19], which has offered 50 activity classes collected from online platforms like YouTube. The activities offered in the dataset include horse riding, pull-ups, diving, running, skipping, etc. Later, the activities are extended into 101 classes with the same human activity category, which is called UCF101. Both datasets offer only RGB data at a resolution of 320 x 240 with a fixed frame rate of 25 frames per second (fps).

The kinetics dataset introduced by [22] consists of a significant dataset for HAR with 700 activity classes with more than 700 video par classes. Each video is captured from YouTube videos lasting for ten seconds. Types of activities included in the data set are human-to-human or human-to-object interactions such as shaking hands, hugging, steering the car, and brushing the floor. NTU RGB+D dataset presented by [20] offers a diverse range of activity classes. Types of action are divided into three categories: eleven mutual activities like pushing, kicking, etc., nine health-related activities such as sneezing, staggering, etc., and 40 daily activities like drinking, reading, etc. It consists of approximately fifty-seven thousand samples in RGB + Depth and in skeleton format. In addition, MSR-Action3D presented by [21] has been a choice for skeleton-based activity recognition. Dataset offers 567 depth map sequences with 20 different hand gesture activity class-like horizontal arm waves, drawing a circle with an arm. Depth maps are captured using a depth camera sensor and are available in 640 x 240 resolution of recorded sequences

Though many datasets offer a diverse range of activity classes to facilitate HAR processes, most of them are related to daily life or health-related activities. From an industrial HAR point of view, there is a lack of a dataset that offers industrial activities, which is further addressed by [10] and presented *InHard* dataset (Industrial Human Activity Recognition Dataset). *InHard* demonstrates the actual industrial activity in an industrial environment, and the dataset is publicly released to facilitate the research progress in the field. The dataset provides various industrial assembly activities in the skeleton and RGB video format to facilitate HAR in the industrial environment. Moreover, it has not been well evaluated by the scientific community.

3. Methodology

A Panasonic 4K camcorder was employed in our experiment to obtain the video from the right side at 45 degrees. Similarly, the Xsens Awinda IMU system is used to capture the joint motions for comparison to the *InHard* (c. [10]) dataset. However, both datasets comprise different settings such as frame rate, skeleton joint numbers, and motion capturing systems. Therefore, we must resample the time to make a consistent frame rate, re-arrange the data structure and retarget the skeleton before comparing.

The *InHard* dataset has offered an actual use case of industrial activities such as assembly of a part, picking components, measuring components, and representing actual industrial setup. The dataset comprises RGB video and IMU data for different participants (person) and in an adequate quantity. The participant's task in the *InHard* dataset is to assemble a component following instruction sets with the help of the UR10 robotic arm, using screws and hooks and a tool such as a screwdriver (c. [10,23]). The same activity with a different job (e.g., assembly of gear components) is proposed to reproduce *InHard* activities. The RGB video from the top view is used to analyze the spatial activities, while the RGB camera from the side helps to acquire (c. Figure 1)

(a) (b) (c) (d) (e)

Figure 1 Validation and comparison of the InHard and newly captured datasets in the lab environment for similar activities but different sensing systems; InHard dataset (a) IMU skeleton, (b) Open pose overlay, (c. [7]), New dataset (c) IMU skeleton, (d) Open pose overlay, (e) RGB spatial.

The methods for HAR based on video and skeleton dataset is described in three categories in subsection 3.1, and 3.2.

3.1 Human activity definition and dataset curation

Before the pre-processing task, it is necessary to explore the dataset to remove any unwanted data. Only an adequate and equal number of samples are provided for the deep learning method for training. Following is the list of nine low-level activity classes of both skeleton and RGB data with several samples in each class.

Table 1 Number of samples in each activities class.

Activity Classes	Assemble System	No Action	Picking Front	Picking Left	Put Down Component	Put Down Screwdriver	Take Component	Take Screwdriver	Turn Sheets
No. of samples	1378	500	456	641	385	461	485	420	224

It is essential to provide equal training data for each activity class to ensure proper learning of the network and the overall accuracy of the network. As deep learning algorithms require a large amount of data for good performance, we have only considered more than 200 samples classes. To facilitate HRC in the industry, the *InHard* dataset community has generalized assembly activities and presented activities that are used in many industrial assembly processes. Activities are divided into low-level activities, consisting of nine action classes (see Table 1), and high-level activities, which comprise 72 detailed action classes for more accurate activity detection.

3.2 HAR modeling based on sequential and temporal memory networks

A fusion of CNN and LSTM (see Figure 2) is proposed to take advantage of both networks as CNN handles spatial information and LSTM takes care of temporal data information. Video frames are given as input to CNN using a pre-trained network inception V3. The inception V3 model is one of the commonly used computer vision techniques for tasks such as object detection. It is pre-trained on the ImageNet dataset consisting of one thousand categories. Inception V3 architecture is built with symmetric and asymmetric blocks. The block includes a series of smaller convolutions, average pooling, and max pooling for faster training and processing of image data [24]. The last output layer of the CNN network is removed to obtain the feature vector. Then this feature vector becomes an input to the LSTM to learn temporal dependency and give the final classification.

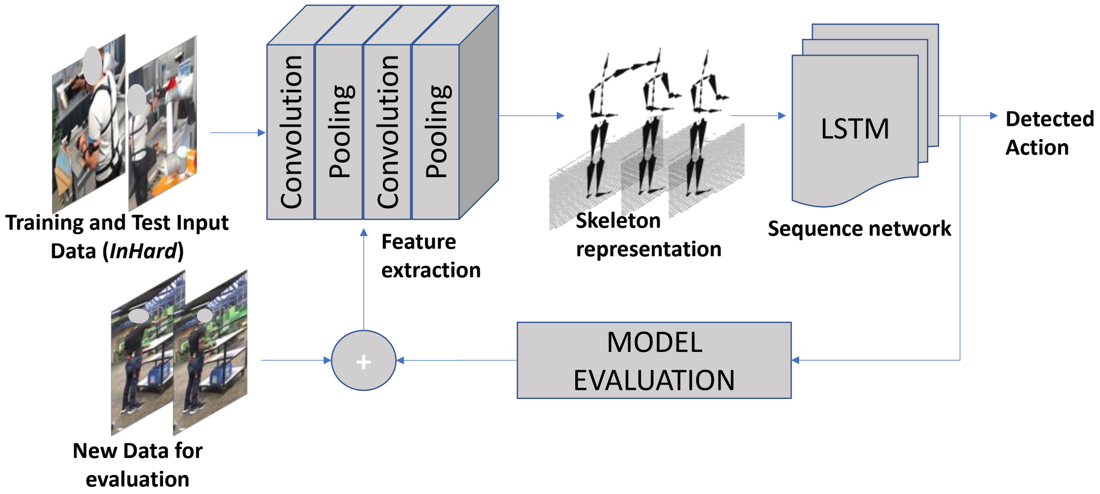


Figure 2 Fusion of CNN and LSTM architecture for action recognition and model evaluation using InHard and new dataset.

Table 2 System configuration for deep learning model

Feature	**Description**
Hardware configuration	All the deep learning model pre-processing and training are done using a high-performance computing (HPC) cluster, which has a specification of AMD 3.35 GHz CPU, NVIDIA Tesla V100 GPU, and 128 GB of RAM.
Data preparation	A model is trained for 14 low-level activity classes with approx. 400 samples per class. 77% of data is used for training, 3% for training validation and the remaining 20% of data is used to evaluate the trained model.
Optimizer	Adamax

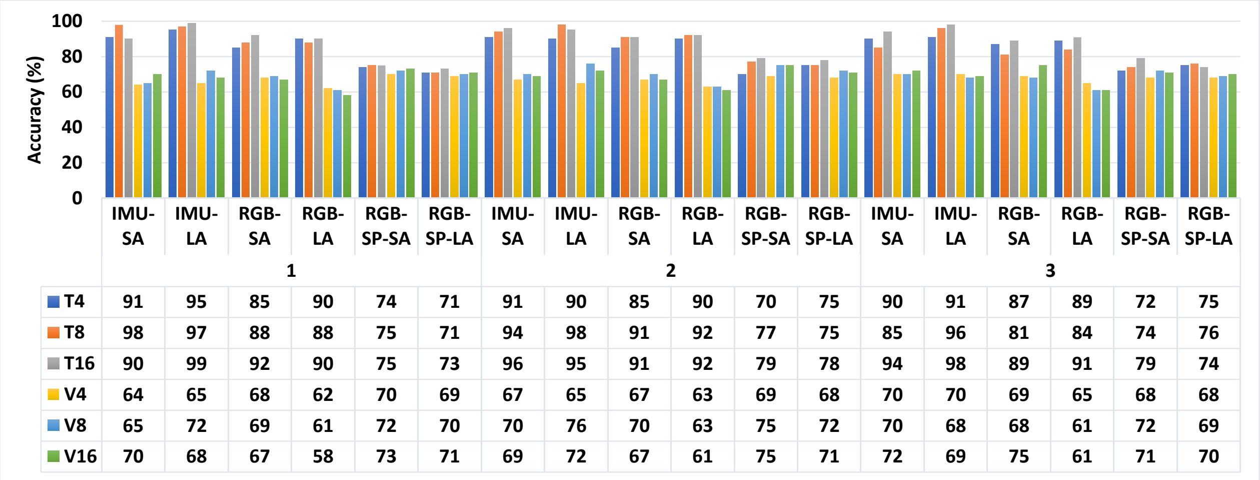
Performance of the *InHard* dataset is tested using a deep learning model: Long Short-Term Memory (LSTM) and Convolution Neural Network (CNN). The accuracy of the dataset is tested using segmented activity recognition using both Skeleton and RGB data. The system configuration details are shown in Table 2. The implementation comprises IMU-based joint data and RGB-based skeleton pose data to further predict and detect human activity in the same time domain. Pose detection and LSTM technique are applied to RGB video data to extract the skeleton pose from video using an open pose library [14]. LSTM processes the extracted skeleton pose data and classifies the activity. LSTM is used to process the temporal dependency of extracted features and classify the activity at the end. We have used a single LSTM model to train skeleton data (BVH files); for RGB data (.mp4 files) training, CNN and LSTM are fused in which CNN is used to extract the spatial feature (resolution) of video frames.

As we are dealing with activity recognition tasks that may include different lengths, we have employed an LSTM model as a final output model for all cases. It is necessary to note that LSTM takes temporal features as an input in which the time-stamp indicates the length of the activities. Some of the *InHard* dataset activities are of different lengths, and this required us to perform training of the model using two different lengths of activity for each method. Accordingly, the activity length is categorized as short and long length activity in which for short length activity, the first 30 seconds of data is considered, and for prolonged length activity, the first 60 seconds of data is considered for each case.

Before starting the training of deep learning models, it is necessary to clean and pre-process the training and test data. The cleaning and pre-processing of the *InHard* dataset include dividing the data frame into smaller batches, extracting skeleton data, and labeling its class. The training process is similar for both Skeleton (BVH) and video-skeleton cases. In both, hierarchical skeleton data is converted into vectors along with its class label. Besides, the data frame is divided into smaller batches depending on the length of the activity (short and long). Then these batches are used to train the LSTM model. The human activity recognition network code is publicly available [25]. The Tensorflow and Keras framework has been implemented for the training deep-learning model. Tensorflow is an open-source platform that provides various machine learning libraries, and Keras is a user-friendly high-level API that runs on top of Tensorflow [26].

4. Result

Based on systematically selected Tensorflow and Keras framework parameters, the early patience is to 4, and the learning rate is set to 0.02. The loss function is set to *categorical cross-entropy* as it is the default choice for classification.



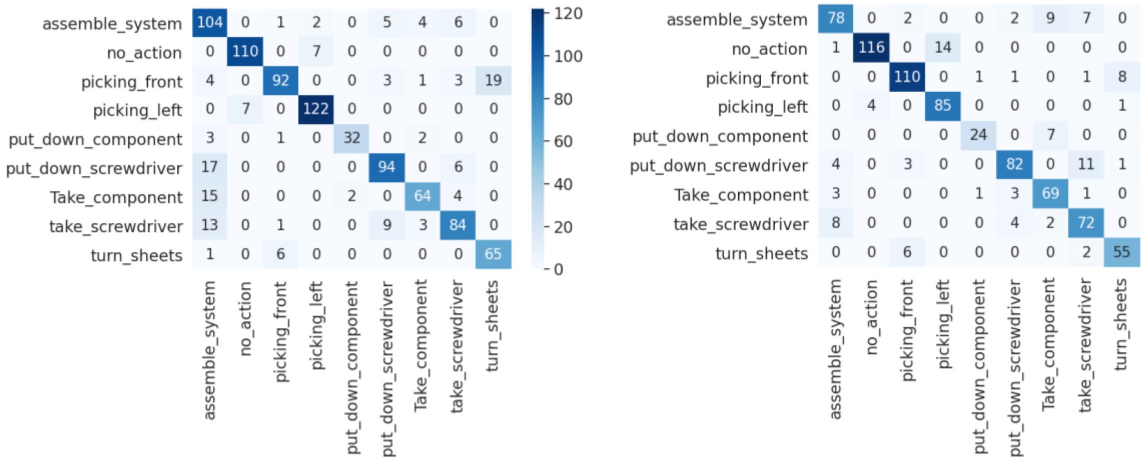
	IMU-SA	IMU-LA	RGB-SA	RGB-LA	RGB-SP-SA	RGB-SP-LA	IMU-SA	IMU-LA	RGB-SA	RGB-LA	RGB-SP-SA	RGB-SP-LA	IMU-SA	IMU-LA	RGB-SA	RGB-LA	RGB-SP-SA	RGB-SP-LA
	1						2						3					
🟦 T4	91	95	85	90	74	71	91	90	85	90	70	75	90	91	87	89	72	75
🟧 T8	98	97	88	88	75	71	94	98	91	92	77	75	85	96	81	84	74	76
⬜ T16	90	99	92	90	75	73	96	95	91	92	79	78	94	98	89	91	79	74
🟨 V4	64	65	68	62	70	69	67	65	67	63	69	68	70	70	69	65	68	68
🟦 V8	65	72	69	61	72	70	70	76	70	63	75	72	70	68	68	61	72	69
🟩 V16	70	68	67	58	73	71	69	72	67	61	75	71	72	69	75	61	71	70

Figure 3 Comparison of model parameters for short activity (SA) or long activity (LA) for training (T) and validation (V) phase.

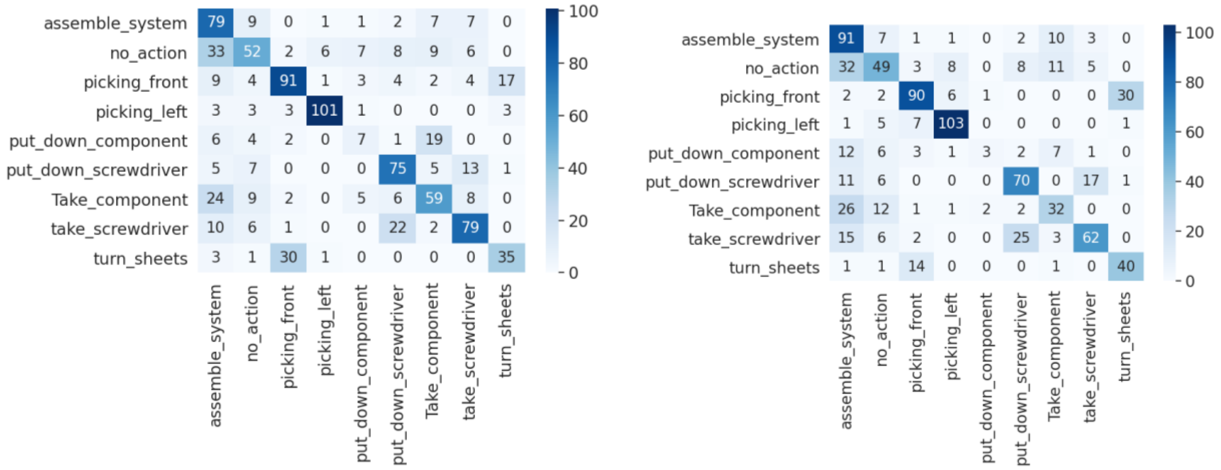
For tuning the models, we have altered different parameters such as optimizer, size of input batches, and the number of hidden layers to obtain optimal parameters. The tuning parameters are epoch: 100, optimizer: 'Adam' and 'Adamax', number of hidden layers: 1, 2, and 3, number of batches: 4, 8, and 16, and the same training process is repeated for two activities length data, i.e., short and long (see Figure 3).

Training result is presented for skeleton (BVH) and video-skeleton data. For skeleton (BVH) and video-skeleton data, the LSTM model is used and trained using a different model structure and parameters for short and long activity length types. The same training procedure is applied for RGB video data on CNN and LSTM network fusion. Training and validation accuracy for each model structure, parameters, and activity length is compared in Figure 3. The result shows the percentage of accuracy for RGB-based skeleton short activity and prolonged activity (RGB-SA and RGB-LA), RGB-based spatial video for short and long activity

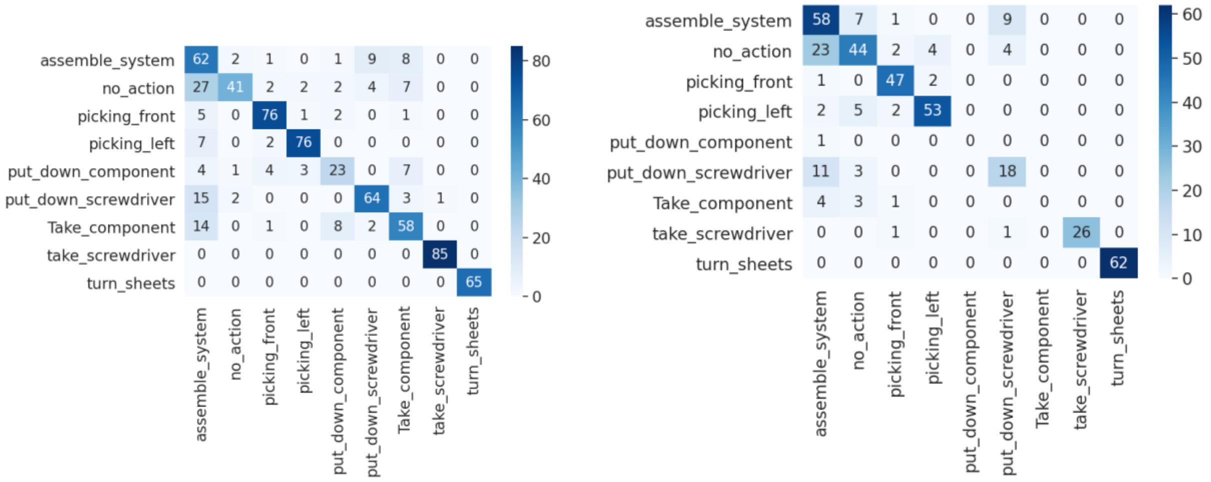
(RGB-SP-SA and RGB-SP-LA), and IMU-based skeleton for short and long activity (IMU-SA and IMU-LA) data for different numbers of hidden layers and batch sizes (4, 8, and 16).



(a) Short activity from IMU skeleton (IMU-SA) (b) Long activity from IMU skeleton (IMU-LA)



(c) Short activity from RGB skeleton (RGB-SA) (d) Long activity from RGB skeleton (RGB-LA)



(e) Short activity from RGB spatial (RGB-SP-SA) (f) Long activity from RGB spatial (RGB-SP-LA)

Figure 4 Confusion matrix representing the comparison of the predicted and detected actions.

The quantitative results based on Figure 3 show that short-activity recognition based on *RGB video skeleton* has achieved model training accuracy of 89% and validation accuracy of 75%. The training accuracy remains 89% for the long-activity, while the validation accuracy drops to 65%. Using the model for evaluation on a test dataset yields 62% and 63% accuracy for short and long activity, respectively. Considering the *IMU skeleton,* the highest training accuracy and validation accuracy of short-activity are 94% and 72%, respectively. The long-activity is 98% for training accuracy and 76% for validation. Using the training weights of the model with the highest training accuracy and validation accuracy, the evaluation with test data yields evaluation accuracy of 85% for short-length activity and 88% for long-activity. For RGB video, a fusion of CNN and LSTM achieved 81% and 78% training and validation accuracy for short activity length and 80% and 72% training and validation accuracy for long activity length. Fusion of model has shown a 77% and 74% evaluation accuracy on test data.

A confusion matrix, which has a size of *n x n*, where *n* is the number of activities, is used to evaluate how accurately the model can classify the activities (c.[27]). The matrix compares the actual activities with the predicted activities (see Figure 4).

5. Discussion

To facilitate HRC, we have tested the *InHARD* dataset for industrial activity recognition using deep learning techniques for two modes of data: IMU skeleton data, and video-skeleton data, among which skeleton activity recognition has shown 88% evaluation accuracy, RGB video model gave 77%, and video-skeleton data is with 63% of evaluation accuracy on *InHard* dataset.

The model's prediction is better when activity length is long because it has acquired more data frame (i.e., the batch size consists of 60 seconds of activity) than in short-length activity where the duration is 30 seconds. Skeleton data provides detailed information about each human pose, more training data, and improved accuracy. Thus, it has been possible to distinguish the classification of activities having slight differences, such as putting down the screwdriver and picking up the screwdriver accurately.

The video-skeleton method using open pose techniques shows poor results compared to the IMU methods. It classifies the activities; however, it gets confused between similar activities such as *Take the screwdriver* and *Put Down Screwdriver*. As the open pose technique highly depends on the person's view for detecting key points on the body for mapping to the skeleton, considering a camera position in a proper orientation with minimum occlusion possibility is crucial for obtaining better results. With the implemented open pose technique, only the required region of interest is considered from the RGB video to avoid noise that may affect the model accuracy. The comparison results with different datasets captured in the lab environment shows less than fifty percent success ratio from the captured twenty operations, while the IMU skeleton accuracy is 67%. On the other hand, the RGB video model (CNN+LSTM) has shown 60% accurate detection for some activities captured in the lab with Panasonic 4K camera. The accuracy evaluation of the proposed methods is still below the accuracy of the human activity recognition that has been published in [17], which employs region-based joint configuration. Reproducible workplace setup does not necessarily yield the same output for reasons such as motion capturing systems, body size variation, and implementation complexities.

Overall results show that human activity recognition for industrial setup is still challenging to detect activities when robots are considered in the loop accurately. Due to the skeleton, body size, capturing system, and model parameters, repeated activities performed in different workplace settings are not straightforward to reproduce. Open source datasets such as *InHard* are helpful to investigate optimal settings for motion capturing and modeling, allowing to exploit the opportunities and identify the inherent challenges regarding activity prediction and detection techniques. However, more datasets must be employed before generalizing human activities and actions detections. This may facilitate the path toward sustainable human and robot collaboration.

6. Conclusion and future outlooks

Human activity recognition in the cognitive production system may change the way humans and machines interact and cooperate for completing tasks. Gathering sufficient data that helps to extensively evaluate the performance and limitations of existing methods is still challenging. The main reasons discussed are model accuracy, data validity, and activity duration. Employing multi-systems for human motion data acquisition such as IMU and Optical cameras, methods such as CNN+LSTM approaches are evaluated for their accuracy. The overall result shows open research questions regarding motion capturing methods, feature mapping, and labeling. Nevertheless, the proposed approach has the potential to improve the way robots learn human motion behavior as co-partners. Future works will address real-time activity recognition with an extended cognitive capability in human-centric workplaces.

Acknowledgments

The authors acknowledge the European Regional Development Fund (EFRE) within the project SMAPS (grant number: 0200545) for the financial support. Similarly, the authors would like to acknowledge the ZIMT department for *Scientific Computing* services at the Universität Siegen.

References

[1] Li S, Wang R, Zheng P, Wang L. Towards proactive human–robot collaboration: A foreseeable cognitive manufacturing paradigm. J Manuf Syst 2021;60:547–52. https://doi.org/10.1016/j.jmsy.2021.07.017.

[2] Pichler A, Akkaladevi SC, Ikeda M, Hofmann M, Plasch M, Wögerer C, et al. Towards Shared Autonomy for Robotic Tasks in Manufacturing. Procedia Manufacturing 2017;11:72–82. https://doi.org/10.1016/j.promfg.2017.07.139.

[3] Ji Z, Liu Q, Xu W, Liu Z, Yao B, Xiong B, et al. Towards Shared Autonomy Framework for Human-Aware Motion Planning in Industrial Human-Robot Collaboration. 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), 2020, p. 411–7. https://doi.org/10.1109/CASE48305.2020.9217003.

[4] Vysocky A, Novak P. Human - Robot collaboration in industry. MM Science Journal 2016;2016:903–6. https://doi.org/10.17973/MMSJ.2016_06_201611.

[5] Kumar S, Savur C, Sahin F. Survey of Human–Robot Collaboration in Industrial Settings: Awareness, Intelligence, and Compliance. IEEE Transactions on Systems, Man, and Cybernetics: Systems 2021;51:280–97. https://doi.org/10.1109/TSMC.2020.3041231.

[6] Bauer A, Wollherr D, Buss M. Human–robot collaboration: a survey. Int J Human Robot 2008;05:47–66. https://doi.org/10.1142/S0219843608001303.

[7] Hartmann B. Human worker activity recognition in industrial environments. KIT Scientific Publishing; 2011. https://doi.org/10.5445/KSP/1000022235.

[8] Roitberg A, Somani N, Perzylo A, Rickert M, Knoll A. Multimodal Human Activity Recognition for Industrial Manufacturing Processes in Robotic Workcells. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, Seattle Washington USA: ACM; 2015, p. 259–66. https://doi.org/10.1145/2818346.2820738.

[9] Voronin V, Zhdanova M, Semenishchev E, Zelenskii A, Cen Y, Agaian S. Action recognition for the robotics and manufacturing automation using 3-D binary micro-block difference. Int J Adv Manuf Technol 2021. https://doi.org/10.1007/s00170-021-07613-2.

[10] DALLEL M, HAVARD V, BAUDRY D, SAVATIER X. InHARD - Industrial Human Action Recognition Dataset in the Context of Industrial Collaborative Robotics. 2020 IEEE International Conference on Human-Machine Systems (ICHMS), 2020, p. 1–6. https://doi.org/10.1109/ICHMS49158.2020.9209531.

[11] Cho NJ, Lee SH, Suh IH. Modeling and evaluating Gaussian mixture model based on motion granularity. Intel Serv Robotics 2016;9:123–39. https://doi.org/10.1007/s11370-015-0190-1.

[12] Ji S, Xu W, Yang M, Yu K. 3D Convolutional Neural Networks for Human Action Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence 2013;35:221–31. https://doi.org/10.1109/TPAMI.2012.59.

[13] Wang P, Liu H, Wang L, Gao RX. Deep learning-based human motion recognition for predictive context-aware human-robot collaboration. CIRP Annals 2018;67:17–20. https://doi.org/10.1016/j.cirp.2018.04.066.

[14] Cao Z, Hidalgo G, Simon T, Wei S-E, Sheikh Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence 2021;43:172–86. https://doi.org/10.1109/TPAMI.2019.2929257.

[15] Shi X, Li Y, Zhou F, Liu L. Human Activity Recognition Based on Deep Learning Method. 2018 International Conference on Radar (RADAR), 2018, p. 1–5. https://doi.org/10.1109/RADAR.2018.8557335.

[16] Liu H, Qu D, Xu F, Zou F, Song J, Jia K. A Human-Robot Collaboration Framework Based on Human Motion Prediction and Task Model in Virtual Environment. 2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), 2019, p. 1044–9. https://doi.org/10.1109/CYBER46603.2019.9066603.

[17] Manns M, Tuli TB, Schreiber F. Identifying human intention during assembly operations using wearable motion capturing systems including eye focus. Procedia CIRP 2021;104:924–9. https://doi.org/10.1016/j.procir.2021.11.155.

[18] Kuehne H, Jhuang H, Garrote E, Poggio T, Serre T. HMDB: A large video database for human motion recognition. 2011 International Conference on Computer Vision, 2011, p. 2556–63. https://doi.org/10.1109/ICCV.2011.6126543.

[19] Reddy KK, Shah M. Recognizing 50 human action categories of web videos. Machine Vision and Applications 2013;24:971–81. https://doi.org/10.1007/s00138-012-0450-4.

[20] Shahroudy A, Liu J, Ng T-T, Wang G. NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE; 2016, p. 1010–9. https://doi.org/10.1109/CVPR.2016.115.

[21] Li W, Zhang Z, Liu Z. Action recognition based on a bag of 3D points. 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, USA: IEEE; 2010, p. 9–14. https://doi.org/10.1109/CVPRW.2010.5543273.

[22] Smaira L, Carreira J, Noland E, Clancy E, Wu A, Zisserman A. A Short Note on the Kinetics-700-2020 Human Action Dataset. ArXiv:201010864 [Cs] 2020.

[23] DALLEL M, HAVARD V, BAUDRY D, SAVATIER X. InHARD - Industrial Human Action Recognition Dataset in the Context of Industrial Collaborative Robotics 2020. https://doi.org/10.5281/zenodo.4003541.

[24] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p. 2818–26. https://doi.org/10.1109/CVPR.2016.308.

[25] Tuli TB, Patel VM, Manns M. HARNets: Human Activity Recognition Networks Based on Python Programming Language. 2022. https://doi.org/10.5281/zenodo.6366665.

[26] Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, et al. TensorFlow, Large-scale machine learning on heterogeneous systems. 2015. https://doi.org/10.5281/zenodo.4724125.

[27] Powers DMW. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation, 2008.

Biography

Tadele Belay Tuli is a researcher and Ph.D. candidate at the University of Siegen, Germany. He received his M.Sc. in Mechatronics Engineering, robotics curriculum from the University of Trento, in Italy. His research interest covers human motion behavior modeling for human-robot collaboration, human activity recognition, and path planning for 3D printing.

Valay Mukesh Patel is an M.Sc. student at the University of Siegen, Germany, in Mechatronics Engineering. His research interest is automation and human activity recognition.

Martin Manns has been head of the Chair of Production Automation and Assembly (FAMS) at the University of Siegen since 2016. Before his appointment, Univ.-Prof. Dr.-Ing. Martin Manns has worked in production research at Daimler AG (2009-2016) and Henkel KGaA (2007-2009) and as a post doctorate fellow at the University of Winsor, Ontario, Canada (2006-2007).

3rd Conference on Production Systems and Logistics

The Contribution Of New Production Technologies And Circular Economy Towards Meeting The Future Demand Of Proton-exchange membrane Fuel Cells – A Literature Review

Patrick Alexander Schmidt[1]

[1]*Fraunhofer Institute for Machine Tools and Forming Technology IWU, Reichenhainer Straße 88, 09126 Chemnitz, Germany*

Abstract

The energy and mobility sectors contribute significantly towards the global CO_2 emissions. The proton-exchange membrane fuel cell finds application in both sectors and represents a possible green and sustainable technology for electricity generation. Current production rates do not satisfy the predicted demand for proton-exchange membrane fuel cells as the diffusion of this technology keeps increasing. Nor does the per-part cost guarantee a globally sufficiently broad application. The industry must overcome technological and economic obstacles to enable higher production rates at a lower cost per unit.

This research gives an overview of current proton-exchange membrane fuel cell production and stacking technologies and provides an outlook on processes that need to be improved to enable faster and lower-cost production. Additionally, the impact of remanufacturing as an end of life option on the circular economy, production, and ecological impact of proton-exchange membrane fuel cells is examined.

The knowledge generated by this research shall support increasing proton-exchange membrane fuel cell production rates to catch up with the predicted demand. Since current research on proton-exchange membrane fuel cell remanufacturing is rare, findings on this topic will support the industry in preparing for circular production processes in the future. Results of the present work include an overview of the current state of production for proton-exchange membrane fuel cells, the areas that need improvement, and the role of a circular economy.

Keywords

Circular economy; End of life; High-rate production; Proton-exchange membrane fuel cell (PEMFC); Remanufacturing.

1. Introduction

With the Paris Climate Agreement, 194 countries and the European Union set themselves to limit global warming due to increasing greenhouse gas emissions to well below 2°C and ideally to below 1.5°C [1]. Germany was responsible for 1.9% of global CO2 emissions in 2019 [2]. In 2019, Germany caused 39% of its CO2 emissions by electricity generation, 23.5% by industry, 21.5% by transport, and 16% by the building sector (only includes heat production, other sources count towards electricity consumption) [3]. In order to counteract the climate crisis, the areas mentioned above need a transformation. One possible technology for reducing CO2 emissions in the transport, industry, and building sectors are proton-exchange membrane fuel cells (PEMFC) combined with green hydrogen. The use of PEMFCs can replace the use of fossil fuels in

DOI: https://doi.org/10.15488/12172

transport and, when installed in buildings as combined heat and power units, generate electricity and at the same time use waste heat efficiently.

Due to limited distribution, the current annual production volume of PEMFC is still shallow. However, there are signals from politics, such as the National Hydrogen Strategy [4], and the economy, which point to an enormously increasing sales market for PEMFC. The Hydrogen Council forecasts that 10 to 15 million cars and 500,000 trucks worldwide will run on fuel cell technology in 2030, and 400 million cars and 15 to 20 million trucks in 2050 [5]. The Japanese car manufacturer Toyota states that its fuel cell cars production will increase tenfold from 3,000 vehicles per year in 2021 [6]. Optimistic forecasts see growth to up to 1.8 million hydrogen vehicles in Germany in 2030. The use of PEMFC is not limited to passenger cars. The technology benefits public transportation with regional trains [7].

In order to meet the anticipated demand for fuel cells, a high-rate production is required (explained in section 3.2). Furthermore, a holistic consideration of the life cycle of the PEMFC is necessary to exploit the sustainability potential fully. Therefore, in addition to manufacturing, it is necessary to examine the aspects of disassembly and recycling of the PEMFC. This paper aims to overview the existing processes for assembly, high-rate production, disassembly, and end of life (EOL) options of the PEMFC and derive necessary developments. The following section deals with the mode of operation, the components, and the application areas of the PEMFC. Subsequently, existing stacking concepts for the PEMFC, possibilities of a high-rate production, and effects on the production costs will be examined. Finally, possibilities and obstacles in the dismantling and recycling process are presented.

2. Operating principle

Fuel cells convert chemical energy from gases or liquids into electrical energy and release heat during this process [8]. In the PEMFC, this process occurs through the oxidation of a fuel and the reduction of an oxidant [9]. The fuel is hydrogen, and the oxidant is atmospheric oxygen. The membrane electrode assembly (MEA) separates the hydrogen and oxygen streams. The MEA consists of a semi-permeable membrane that carries an electrolyte and two electrodes with a catalyst layer [10]. The separating membrane is only permeable to the hydrogen protons produced during oxidation on the anode side of the catalyst but not to the electrons. It is therefore referred to as a proton exchange membrane (PEM). The flow of electrons takes a diversion via an external circuit to the cathode and generates electricity in the process. The electrons are then absorbed by the atmospheric oxygen at the cathode, forming water.

In addition to the PEMFC, there are other types of fuel cells. These differ in the electrolytes used, the fuel for the anode, the level of power, the operating temperature, and the electrical efficiency, resulting in different areas of application [9].

The application areas of the PEMFC can be divided into stationary, portable, and mobile applications. Applications in the portable sector include a power supply for small consumers. Stationary power generation up to the megawatt range and combined heat and power generation are other application areas [11]. Mobile uses for the PEMFC are trucks, submarines, cars, and buses [9]. The PEMFC is well suited for mobile applications, as the power output can be controlled dynamically [12]. This study focuses on the PEMFC.

3. Assembly

3.1 Stacking concepts

The voltage of one fuel cell amounts to 1V approximately. Most applications, however, require higher voltage. Therefore, several individual cells, consisting of alternately stacked bipolar plates (BPP) and MEA, are combined into a stack. This process can be performed manually or automated and must be prepared in previous work steps further explained in the following [13].

The MEA is the most crucial fuel cell component and consists of two components, which are the catalyst-coated membrane and the gas diffusion layer. The decal process is a method to produce the catalyst-coated membrane [14]. In this process, the catalyst material (a mixture of carbon substrate and platinum) is first applied to a carrier medium (decal) and transferred to the polymer membrane in a further step by hot pressing. To produce the gas diffusion layer carbon paper or fabric is needed. It is made from cut carbon fibres and subsequently impregnated and graphitised. A microporous layer is subsequently applied to increase the water balance of the electrode. The finished gas diffusion layer is then glued to both sides of the catalyst-coated membrane by hot pressing and cut into shape as a MEA. The BPPs are still missing for a complete fuel cell. They can be made from coated metal or a graphite composite material. While the composite plate can already be provided with a seal after the manufacturing process, the metal variant requires further work steps beforehand. The half-plates must first be cut out, which is usually done by laser, and then two half-plates must be welded on top of each other to form a complete BPP. A gasket is applied to the combined plate after a leak test. The gasket serves as insulation between BPP and MEA. This concludes the preparation for the stacking.

The stacking process starts with the lower endplate and the lower current collector. A BPP with a gasket and a MEA are stacked on top of each other alternating. This process is repeated until the desired number of fuel cells has been reached. Then, the upper current collector and the upper endplate follow, in which the media supply lines are also located. The individual components must be placed as precisely as possible. In the next step, the complete stack is pressed together to connect everything tightly and minimise the resistance of the contact surfaces. It is essential to ensure that the pressing force is applied evenly everywhere during this process. If this is not the case, significant reductions in performance and service life cannot be ruled out. The compressed state must be guaranteed permanently. The stack is held together by either tension bands or tension rods. Subsequently, an initial leak test takes place, such as pressure drop and flow tests. Before the stack is placed in a housing and covered by the distributor plate, final work steps are necessary, such as attaching the contacts and current busbars and the cell voltage monitoring module for the cell voltage. The so-called running-in on the test bench follows, where all mediums (hydrogen and oxygen) are connected. Through this process, the stack's performance can be determined. Finally, another leak test may be carried out, after which the fuel cell is installed with other system components and is ready for use [14].

3.2 High-rate production

The stacking process described in the previous section can be carried out in various ways. While the stacks are still partly stacked by hand, partially or fully automated processes have been developed in the meantime. Automated processes use carousel devices or conveyor and feed systems, although manual work steps are still necessary in some cases. Pick-and-place robots could automate these processes [15]. The extended use of robots would be a first step in the direction of high-rate production. A crucial development, especially regarding the capacities, is required in the future [16].

In the context of this paper, the term "high-rate production" concerning the fuel cell means an approximation to the production figures from the automotive industry. The cycle time for a vehicle is 60 seconds on average. In order to adapt the fuel cell stack production to this, ten cells must be stacked in one second, based on a stack with 400 to 600 cells. That situation, however, assumes that 100% of cars plant’s production need a fuel cell stack and is therefore relevant for highly specialised fuel cell production plants. If a supplier wanted to produce, for instance, 20.000 fuel cell stacks a year with 400 cells per stack, the cycle time would be roughly about 2.5s per fuel cell (assumptions: 240 days and 24h of work).

Nevertheless, the cycle times are still insufficient for the envisaged demand, so parallelisation and modularisation of the production capacities are necessary. A modular approach might allow to adapt steadily to increasing demands. Irrespective of this, production volumes must be increased to meet the planned future sales volumes of fuel cell electric vehicles (FCEVs), and new and faster production facilities must be found as a result. The necessary production capacities entail several prerequisites that need consideration as early as the planning stage of the plants. The quality and the safety of the fuel cell must not be reduced by shorter cycle times. Precision and repeatability are essential, and the handling of sensitive components must also be guaranteed in terms of quality, even in automated production. Some of the components, such as the MEA and gaskets, are flexible and thin, complicating the automated handling even further [17]. Since the same components from different suppliers can differ in shape, size and tolerances, machines must be adaptable. The same applies to the produced stacks, which have different dimensions depending on performance. As a result, it is advantageous if machines already have a high degree of adaptability, which prevents long conversion times or even breakdowns. For a smooth production process, bottlenecks in the components must be taken into account and calculated in advance. Subsequently, the cycle and throughput times must be adaptable [16].

In principle, however, making only the stacking process faster does not suffice to reach the desired production capacities. The individual component production must be improved, too. The production of BPPs and MEAs and the connection of the two components has great potential. The components could be produced in the roll-to-roll process in the future. In this process, the material is unwound from a roll, processed, and then rolled up again, which speeds up further processing and makes it possible to produce complete cells in a short time and separate them. Hence, they are directly available for the production of stacks [18]. However, even if the individual components are modified, it is necessary to observe several criteria to ensure that the fuel cell meets all requirements. These requirements are mainly regarding the performance of the fuel cell stack and system, where, among other things, thermal and electrical resistance must be minimised, but the conductivities must be equally maximised. Mediums such as hydrogen and oxygen must also be supplied and discharged without leakage, particularly when ambient temperatures change.

Furthermore, low masses and dimensions are advantageous. In addition to these aspects, manufacturing properties and environmental influences are also relevant, in connection with which material selection, manufacturing processes, and recycling possibilities are to be mentioned [19]. These criteria are described in more detail in section 4. Currently, there are still some hurdles to overcome to fully automated production. New processes with a direct coating of the electrode onto the membrane or gas diffusion layer are necessary. The production time can be significantly reduced with a continuous lamination process. Faster and, above all, defect-free production methods must also be planned for BPP, regardless of the material. In general, the quality factor plays an essential role in both the components and the final products, which is why the final controls or the methods for checking them must also be improved and accelerated [20].

Academia and industry are aware of the necessity to increase stacking velocity and stacking scalability. Research is performed on robotised high-speed stacking processes that aim to increase stacking rates through gripping several layers of cells at once or through the inclusion of in-line quality control to further reduce cycle times [21,22]. Researcher designed an automated workstation layout for fuel cell stacking through the parallelisation and decoupling of process steps [23]. Stacking velocities of approx. 2s per stack was reached. An outlook was given that cycle times below 2s are possible but would require high development costs. An increase in production capacities also has a positive effect on the production costs, as explained in more detail in the next section.

3.3 Cost factors

Even though the costs for PEMFC systems have fallen by about 50% in the last 15 years, they are still too high to achieve a breakthrough for the PEMFC as an alternative on the vehicle market alongside battery-powered electric vehicles. The high costs are expected to decrease in the future due to increased demand and the associated higher production volumes [24].

In principle, the costs for the complete fuel cell system consist of the fuel cell stack costs and the costs of the system components (for instance, air compressor, water separator, and hydrogen filter). The focus in this work is on the stack and its components since in this area, more components are highly dependent on the absolute production number. For small production numbers (1,000 stacks/year), five components are responsible for over 80% of the costs. If the projected production figures of 500,000 stacks/year are reached, only two components (BPP and catalyst) are responsible for 66% of the costs. Therefore, the total price can be reduced most effectively by saving costs through an increased production volume [16]. According to a report by the U.S. Department of Energy (DOE) from 2018, the price of a fuel cell system per kilowatt net power (kWnet) should be reduced to $30 in the future (based on an 80 kWnet PEMFC and production quantities of 500,000 stacks/year). As an overview, the DOEs findings on the costs of the individual components, depending on the annual production volume, are shown in Table 1.

Table 1: Influence of production volume on the costs of components and the entire stack [24].

Yearly production rate of	1,000 stacks	10,000 stacks	100,000 stacks	500,000 stacks
BPPs	1,554$	486$	404$	388$
MEAs	6,546$	2,320$	1,121$	915$
Other components	1,278$	700$	135$	127$
Cost complete stack	9,533 $	3,504$	1,722$	1,479$
Cost complete stack (per kWnet)	119.16$	43.8$	21.52$	18.49$

However, these prices cannot be achieved by increasing capacity alone, which is why the production methods and the materials used must be changed [18]. For BPP, various materials and the production variants required were investigated in terms of the production numbers and costs required. Hydroforming (internal high-pressure forming) is more cost-effective than progressive stamping for large production volumes when using stainless steel. If several panels are produced simultaneously, fewer assembly lines are required, which directly impacts costs and processing time. Another way to decrease cost by changing the material is using an alternative stainless-steel alloy. A cost reduction of $0.13 to $0.21 per kW net could be reached, while there are no disadvantages in use.

Further savings are also possible by using a different plating process. Even if cost reductions are possible by reducing the platinum content and increasing the power density, prices still depend on the material costs [25]. Especially regarding precious metals such as platinum, the material price must be paid regardless of higher production numbers. For this reason, fluctuations must always be taken into account, and the use of other materials should also be considered if necessary.

4. Disassembly

After a defect, caused, for example, by impurities in the MEA, disassembly of the PEMFC is a prerequisite for reuse or further use. In principle, the process occurs in reverse order to assembly so that peripheral components on the stack are removed first (e.g., cell voltage monitoring unit and cooling fan). In the next step, the bracing is loosened. Then the endplates and the current collectors can be dismantled to reach the individual cells. The cells are lifted off one after the other and can then be disassembled into their components (BPP, gasket, gas diffusion layer and MEA) [26]. However, separating the sensitive parts bears the risk of damaging the components and making them unusable.

Similarly, it is hardly possible to replace individual cells, and inaccurate contact surfaces can lead to significant performance losses. For this reason, the introduction of a conductive intermediate layer between the BPPs was considered, which would also make it easier to detach them from each other. Flexible graphite could be used for this purpose. It also compensates for the poorer surface quality of the plates due to its deformability, thus reducing production costs and ensuring better current transmission. At the same time, different loads can be better compensated, which has a positive effect on the lifetime of the entire stack [27]. A similar patent from the Toyota company involves the use of an adhesive intermediate layer of fluoroplastic or silicone resin that can be detached by heat and serves as a seal. This method can also be used for other types of fuel cells and is also intended to facilitate the dismantling process. For this purpose, heat is applied to the adhesive layer and pressed apart with the help of a wedge, which makes it easier to apply the heat. In order to be able to remove residues of the adhesive better and accelerate the separation process, a heat-dissipating agent is applied along with the adhesive layer, which causes the dissolved adhesive to contract again [28].

With the help of these ideas, a reduction of the disassembly time is possible, which proves to be especially important with increasing numbers of FCEVs in the future. It becomes clear that these aspects must be considered before assembly during the planning phase of the stack. Hence, new design guidelines must be considered and used. "Design for disassembly" (DfD) is an approach that aims to simplify as well as speed up the disassembly process, make it more cost-effective, and recycle as many materials and components as possible. The right choice of materials, connecting elements, and the product's design are fundamental steps in the development. The disassembly process can only be improved with sufficient knowledge about the developed object's structure, its use, and the physical and technical limits of the disassembly process [29]. Another approach is the so-called Design for Remanufacturing (DfRem), where several design guidelines for products are followed to facilitate the EOL phase [30–32]. These guidelines include enabling easy, non-destructive disassembly through accessibility, modularity, ease of cleaning and handling, designing for multiple life cycles, resistance to wear and tear, and considering the EOL phase already during the product development process [29,30,33,34]. The DfRem method is primarily used to improve the possibilities for closed-loop capability, which will be defined in more detail in the next chapter.

5. End of life options

5.1 Proton-exchange membrane fuel cell remanufacturing

For the PEMFC to be justifiably described as a sustainable technology, it is necessary also to consider the end of the product life cycle. It can only be used in a genuinely sustainable manner if a circular economy exists in addition to the use of green hydrogen. After focusing on dismantling as a prerequisite for recycling the PEMFC in the last chapter, the following section looks at various EOL options, i.e., possibilities for returning the PEMFC to the product cycle at the end of its life cycle.

The various EOL options differ in several aspects. A distinction is made whether a product is reused in the same application or with a different purpose [33,35]. Furthermore, there is a difference in the point at which the product is reintroduced into the life cycle [36]. The re-entrance point influences the amount of lost energy and materials [37]. EOL options in ascending order of lost materials and energy are Reuse/repair, remanufacturing (product recycling), recycling (material recycling), energy recovery and landfill disposal. Given the cradle-to-cradle approach no waste should be produced. Thus, the first options mentioned should be preferred, and the latter avoided. Next, remanufacturing as a possible EOL option for the PEMFC will be discussed in more detail.

Remanufacturing, or product recycling, is the reprocessing of a product after its use phase to the quality level of a new product [30]. The exact process steps of remanufacturing are variable depending on the application but can be described as follows: Old part procurement, testing/sorting, cleaning, refurbishment, reassembly and a final test [31,36,37]. Various studies of use cases show that remanufacturing influences costs positively and, in particular, decreases environmental impacts compared to new production [36,38]. Remanufacturing results in less dependence on critical raw materials, advantages for the user due to lower prices and strategic advantages for the manufacturer [36]. Examples of remanufacturing applications in the automotive sector include engines, gearboxes, starters and turbochargers [36]. From the field of business mathematics, there are many publications on remanufacturing [39,40]. However, there is a lack of a deeper consideration of remanufacturing for PEMFC from a production engineering perspective. In this consideration, it is crucial to take a holistic approach where economic, environmental, and technological factors are considered simultaneously to develop the optimal process. Many obstacles exist that give the remanufacturing process its complexity. These include low volumes, uncertainty about the number and condition of EOL parts, increasing product complexity and disassembly as a cost driver due to manual processes and low product know-how [34,36,38,41,42].

5.2 Component recycling

While product recycling for the PEMFC has hardly been practised so far, there are already established processes for recycling at the material level, with which different materials of the PEMFC can be recovered. For this purpose, the stack must be broken down into its components. Peripheral components, such as control electronics, can be recycled via conventional e-waste [13]. Purely metallic components such as tension rods, current collector plates and end plates are further processed via metal scrap [43]. Other specific processes exist for the individual components of the fuel cell, which will be discussed in more detail below.

There is a strong focus on the recycling of MEA [13]. One reason is that fuel cell defects can often be traced back to the MEA. During the fuel cell operation, the formation of pores or the accumulation of impurities from the fuel might damage the membrane in the MEA [43]. Another reason is the material cost. About 42g of platinum is installed per stack, which is why the recovery of the materials has a high financial incentive [44]. Furthermore, the dissemination of FCEVs will lead to a sharp increase in platinum consumption, as the needed amount of platinum is ten times higher than in vehicles with internal combustion engines [44].

Processes for recovering precious metals from vehicle catalytic converters of internal combustion engines cannot be used for the PEMFC, as toxic hydrogen fluorides are produced during the combustion of the electrode [43]. Therefore, the catalyst is recovered by chemical extraction [45]. Compared to the catalyst, the recovery of the membrane has not been in focus so far but is becoming significantly more relevant because of increasing production numbers. The recovery of the membrane is complicated because it merges with the gas diffusion layer and the catalyst layer during the operation of the PEMFC, which is why the economic benefit of recovery still needs to be examined in detail [13].

Recycling of BPP depends on the material used. Metallic BPP can be recycled via ordinary metal scrap [13,43]. Graphic BPP made from thermoplastics can be cleaned and reprocessed into granulate for BPP production by injection moulding. In the case of BPP made from thermosets, re-melting is not possible, so only further use as, e.g. filling material is possible [13].

6. Conclusion

The production of PEMFCs is expected to grow, as they represent a sustainable alternative to existing technologies in several application areas. Various developments are still necessary to ensure that the diffusion of PEMFCs is successful in the future.

The current production rate of PEMFC is not sufficient to meet future demand. Steps towards high-rate manufacturing have already been taken using pick-and-place robots, carousel devices, feeding systems, and direct coating of the electrode by the roll-to-roll process. The mentioned measures, however, are currently not yet sufficient. There is still potential for optimisation in the production of the individual components and the necessary function tests. Parallelisation and decoupling of process steps are necessary to increase production rates. Furthermore, production costs must be reduced in the future, partly because the fuel cell costs make up too large a proportion of FCEVs. As production numbers increase, more cost-effective manufacturing processes can be applied, among other things. Existing EOL options and dismantling processes also need to be improved to increase the recyclability of PEMFCs. Inadequate dismantling processes hamper remanufacturing. These processes constitute a significant cost factor in remanufacturing due to mainly non-uniform and manual processes. If dismantling is considered in earlier life cycle phases, through DfD and DfRem, process costs will decrease. Increased recyclability reduces environmental pollution and dependence on critical raw materials.

Fraunhofer IWU's research efforts will address several of the problems mentioned. For example, there are already concepts concerning the component design of individual components and about high-rate technologies and manufacturing processes for producing these components. The degree of automation can be increased with robot-based manufacturing processes. However, alternative technologies are required concerning the desired cycle times of several cells per second. In this regard, there are already initial industry-oriented solutions that enable such production rates through continuous flow processes. For the further qualification of these processes, it is necessary to implement corresponding test facilities to carry out quality-relevant and cycle-time-specific optimisations using the demanding components of the fuel cells. Other innovative concepts about fuel cell design with adaptive assembly elements intend to increase the efficiency of the fuel cell in operation, especially with fluctuating ambient temperatures. These also need to be further qualified regarding an automated high-rate production.

Furthermore, the economic advantages of remanufacturing the PEMFC as an EOL option are investigated. The necessary steps to improve the EOL processes and cycle capability of the PEMFC will be analysed. Another research topic is the standardisation and automation of the PEMFC disassembly process.

References

[1] United Nations, 2015. Paris Agreement.

[2] Global Carbon Project, 2020. CO2-Emissionen: Größte Länder nach Anteil am weltweiten CO2-Ausstoß im Jahr 2019 [Graph]. In Statista. https://de.statista.com/statistik/daten/studie/179260/umfrage/die-zehn-groessten-c02-emittenten-weltweit/. Accessed 24 September 2021.

[3] Roland Berger GmbH, 2020. Potenziale der Wasserstoff- und Brennstoffzellen-Industrie in Baden-Württemberg.

[4] BMWi, 2020. Nationales Reformprogramm 2020 - Die Nationale Wasserstoffstrategie.

[5] Hydrogen Council, 2017. Hydrogen scaling up: A sustainable pathway for the global energy transition.

[6] Toyota Europe, 2018. Toyota moves to expand mass-production of fuel cell stacks and hydrogen tanks towards ten-fold increase post-2020. Toyota Europe, May 24.

[7] Hebling, C., M. Ragwitz, T. Fleiter, U. Groos, D. Härle, A. Held, M. Jahn, N. Müller, 2019. Eine Wasserstoff-Roadmap für Deutschland.

[8] Lipman, T.E., Weber, A.Z., 2019. Fuel Cells and Hydrogen Production. Springer New York, New York, NY.

[9] Jörissen, L., Garche, J., 2017. Wasserstoff und Brennstoffzelle, in: Töpler, J., Lehmann, J. (Eds.), Wasserstoff und Brennstoffzelle. Technologien und Marktperspektiven, 2., aktualisierte und erweiterte Auflage ed. Springer Vieweg, Berlin.

[10] Kurzweil, P., 2013. Brennstoffzellentechnik. Springer Fachmedien Wiesbaden, Wiesbaden.

[11] Badenhop, T., Schellen, M., 2017. Brennstoffzellen in der Hausenergieversorgung, in: Töpler, J., Lehmann, J. (Eds.), Wasserstoff und Brennstoffzelle. Technologien und Marktperspektiven, 2., aktualisierte und erweiterte Auflage ed. Springer Vieweg, Berlin.

[12] TÜV Süd, 2021. PEM-Brennstoffzelle. https://www.tuvsud.com/de-de/indust-re/wasserstoff-brennstoffzellen-info/brennstoffzellen/pem-brennstoffzelle. Accessed 18 October 2021.

[13] Ahlfs, S., Goudz, A., Streichfuss, M., 2020. Die Brennstoffzelle. Springer Fachmedien Wiesbaden, Wiesbaden.

[14] Kampker, A., Ayvaz, P., Schön, C., Reims, P., Krieger, G., 2020. Produktion von Brennstoffzellenkomponenten.

[15] Kampker, A., Ayvaz, P., Schön, C., Reims, P., Krieger, G., 2020. Produktion von Brennstoffzellensystemen.

[16] Porstmann, Wannemacher, Richter, 2019. Overcoming the Challenges for a Mass Manufacturing Machine for the Assembly of PEMFC Stacks. Machines 7 (4), 66.

[17] Fowler, D., Gurau, V., Cox, D., 2019. Bridging the Gap between Automated Manufacturing of Fuel Cell Components and Robotic Assembly of Fuel Cell Stacks. Energies 12 (19), 3604.

[18] Huya-Kouadio, J.M., James, B.D., Houchins, C., 2018. Meeting Cost and Manufacturing Expectations for Automotive Fuel Cell Bipolar Plates. ECS Trans. 83 (1), 93–109.

[19] Cooper, J.S., 2004. Design analysis of PEMFC bipolar plates considering stack manufacturing and environment impact.

[20] U.S. Department of Energy, 2015. Fuel Cell Technologies Office: Multi-Year Research, Development, and Demonstration Plan. 3.5 Manufacturing R&D. U.S. Department of Energy. https://www.energy.gov/eere/fuelcells/downloads/hydrogen-and-fuel-cell-technologies-office-multi-year-research-development. Accessed 3 February 2021.

[21] Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA, 2022. H2FastCell – Pilotanlage für die Produktion von Brennstoffzellensystemen - Fraunhofer IPA. https://www.ipa.fraunhofer.de/de/referenzprojekte/H2FastCell.html. Accessed 21 January 2022.

[22] wbk - Institut für Produktionstechnik, 2022. Projektbeschreibung EMSigBZ: Entwicklung eines modularen und skalierbaren Produktionssystems zur Herstellung von Brennstoffzellen-Stacks. https://www.wbk.kit.edu/wbkintern/Forschung/Projekte/EMSigBZ/?site=home. Accessed 21 January 2022.

[23] Bobka, P., Gabriel, F., Römer, M., Engbers, T., Willgeroth, M., Dröder, K., 2019. Fast Pick and Place Stacking System for Thin, Limp and Inhomogeneous Fuel Cell Components, in: Wulfsberg, J.P., Hintze, W., Behrens, B.-A. (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 389–399.

[24] Brian D. James, Jennie M. Huya-Kouadio, Cassidy Houchins, Daniel A. DeSantis: Strategic Analysis, Inc. Mass Production Cost Estimation of Direct H2 PEM Fuel Cell Systems for Transportation Applications: 2018 Update.

[25] Thompson, S.T., James, B.D., Huya-Kouadio, J.M., Houchins, C., DeSantis, D.A., Ahluwalia, R., Wilson, A.R., Kleen, G., Papageorgopoulos, D., 2018. Direct hydrogen fuel cell electric vehicle cost analysis: System and high-volume manufacturing description, validation, and outlook. Journal of Power Sources 399, 304–313.

[26] Schiemann, J., Kerßenboom, A., Prause, H.J., Peil, S., 2007. Verwertung von Brennstoffzellen und deren Peripherie-Systeme.

[27] Ruge, M., Schmid, D., Buechi, F., 2003. Verfahren und Vorrichtung zum Stapeln von Brennstoffzellen 2003. Accessed 27 January 2021.

[28] Suzuki, H., Yodoshi, N., Tejima, G., Nakashima, T., Akagawa, R., 2005. Fuel cell disassembly method. Accessed 3 February 2021.

[29] Abuzied, H., Senbel, H., Awad, M., Abbas, A., 2020. A review of advances in design for disassembly with active disassembly applications. Engineering Science and Technology, an International Journal 23 (3), 618–624.
[30] Nasr, N., Thurston, M., 2006. Remanufacturing: A Key Enabler to Sustainable Product Systems.
[31] Freiberger, S., 2005. Design for Recycling and Remanufacturing of Fuel Cells, in: Proceedings / Fourth International Symposium on Environmentally Conscious Design and Inverse Manufacturing. 2005 4th International Symposium on Environmentally Conscious Design and Inverse Manufacturing, Tokyo, Japan. 12-14 Dec. 2005. Union of EcoDesigners, Tokyo, pp. 466–471.
[32] Hesselbach, J., Herrmann, C., 2011. Glocalized Solutions for Sustainability in Manufacturing. Springer Berlin Heidelberg, Berlin, Heidelberg.
[33] VDI, 2002. Recyclingorientierte Produktentwicklung. Accessed 11 June 2021.
[34] Kurilova-Palisaitiene, J., Sundin, E., 2014. Challenges and Opportunities of Lean Remanufacturing. Int. J. Automation Technol. 8 (5), 644–652.
[35] Steinhilper, R., 1999. Produktrecycling: Vielfachnutzen durch Mehrfachnutzung. Fraunhofer-IRB-Verl., Stuttgart.
[36] Lange, U., 2018. Ressourceneffizienz durch Remanufacturing – Industrielle Aufarbeitung von Altteilen.
[37] Sundin, E., 2004. Product and process design for successful remanufacturing. Zugl.: Linköping, Univ., Diss., 2004. Univ, Linköping.
[38] Butzer, S., Schötz, S., Steinhilper, R., 2017. Remanufacturing Process Capability Maturity Model. Procedia Manufacturing 8, 715–722.
[39] Sitcharangsie, S., Ijomah, W., Wong, T.C., 2019. Decision makings in key remanufacturing activities to optimise remanufacturing outcomes: A review. Journal of Cleaner Production 232, 1465–1481.
[40] Liu, C., Zhu, Q., Wei, F., Rao, W., Liu, J., Hu, J., Cai, W., 2019. A review on remanufacturing assembly management and technology. Int J Adv Manuf Technol 105 (11), 4797–4808.
[41] Golinska, P., Kawa, A., 2011. Remanufacturing in automotive industry: Challenges and limitations. JIEM 4 (3).
[42] Freiberger, S., 2007. Prüf- und Diagnosetechnologien zur Refabrikation von mechatronischen Systemen aus Fahrzeugen. Zugl.: Bayreuth, Univ., Diss., 2007. Shaker, Aachen.
[43] Simons, A., Bauer, C., 2015. A life-cycle perspective on automotive fuel cells. Applied Energy 157, 884–896.
[44] Wittstock, R., Pehlken, A., Wark, M., 2016. Challenges in Automotive Fuel Cells Recycling. Recycling 1 (3), 343–364.
[45] Xu, F., Mu, S., Pan, M., 2010. Recycling of membrane electrode assembly of PEMFC by acid processing. International Journal of Hydrogen Energy 35 (7), 2976–2979.

Biography

Patrick Alexander Schmidt (*1993) M. Sc., is a research fellow in the body in white, assembly and disassembly department at the Fraunhofer Institute for Machine Tools and Forming Technology IWU. He has been a research fellow at Fraunhofer IWU since 02.2020. In 2019, he completed his master's degree in industrial engineering with a specialisation in mechanical engineering at Chemnitz University. Previously, he studied industrial engineering at Pforzheim University, with an exchange semester at the Tecnológico de Monterrey, Mexico. His research areas include PEM fuel cells, disassembly processes, circular economy, and remanufacturing.

3rd Conference on Production Systems and Logistics

Overcoming Data Scarcity In The Quality Control Of Safety-Critical Fibre-Reinforced Composites By Means Of Transfer And Curriculum Learning

Florian Brillowski[1], Vanessa Overhage[1], Thorsten Tegetmeyer-Kleine[2], Benjamin Hohnhäuser[3]

[1]Institut für Textiltechnik der RWTH Aachen University, Aachen, Germany
[2]MABRI.VISION GmbH, Aachen, Germany
[3]Gesellschaft zur Förderung angewandter Informatik e.V., Berlin, Germany

Abstract

Fibre-reinforced composites are one promising material class to provide a response to the increasing environmental awareness within society. Due to their excellent lightweight potential, fibre-reinforced composites are preferably employed in safety-critical applications, requiring extensive quality control (QC). However, commercially available QC systems are only able to measure fibre deviations, not directly detecting the error itself. In consequence, a worker is required to perform a manual inspection.

Artificial intelligence and especially convolutional neural networks (CNN) offer the opportunity to directly detect and classify defects. However, to train the corresponding algorithms large amounts of data are required, which are often inaccessible in production. Artificial augmentation of the available data is a popular approach to tackle this problem, yet, resulting most of the time in undesired overfitting of the CNN.

Therefore, in this contribution we examine the transfer of human learning behaviour elements to algorithms in form of transfer learning (TL) and curriculum learning (CL). The overall aim is to research, whether CL and TL are appropriate approaches to address data scarcity in e.g. production environments. Therefore, we perform our research on the error detection of three-dimensional shaped fibre-reinforced textiles.

Keywords

Machine Learning; Quality Control; Data Scarcity; Composites; Curriculum

1. Introduction

In light of increasing environmental awareness and a growing resource responsibility within society, lightweight construction solutions are becoming increasingly important. Besides in sports, these constructions are especially used in transportation or engineering to reduce moved masses and hence lower the pollution caused by emissions. Fibre-reinforced composites in particular are preferably employed in safety-critical lightweight applications due to their excellent mechanical properties in relation to their low weight [1]. As a result, many lives depend on the proper function and reliability of fibre-reinforced composites. To prevent fatal component failure, the quality of the manufactured parts is thoroughly tested during and after the production process. Various automated quality control (QC) systems already exist for this purpose. Yet, commercially available systems only measure deviations in fibre orientations [2]. Subsequently, a worker must determine the specific defect in a manual visual inspection. The inspection

DOI: https://doi.org/10.15488/12191

process requires a high level of concentration, is repetitive and exhausting for the worker, potentially resulting in errors the longer an inspection takes [3].

Convolutional neural networks (CNN) are a promising approach to directly detect varying defects without signs of fatigue. However, a large amount of training data is required for each material and defect. Especially for fibre-reinforced composites, many reinforcing textiles are available. Subsequently, this results in a vast amount of defect-material combinations. In industrial environments, data is often only available to a limited extent or not at all due to a lack of integrated sensors [4]. For this purpose, data augmentation techniques such as mirroring or rotating image data have already been developed but can quickly result in an overfitting of a CNN due to lacking data diversity [5]. As a consequence, there is a demand for a machine learning concept that allows the development of an adaptive QC system that uses limited amounts of data in an efficient way.

Therefore, in this contribution we investigate to what extent concepts comparable to the human learning behaviour (e.g. curricula with increasing complexity) can be transferred to algorithm-based learning. In this context, transfer learning (TL) as well as curriculum learning (CL) are examined for defect detection during three-dimensional shaping of reinforcing textiles. The overall aim is to research, whether CL and TL are appropriate approaches to tackle data scarcity in e.g. production environments. Different CNN architectures are evaluated during hyperparameter optimisation and thereafter the results of TL and CL are compared to a regular (vanilla) training approach. In conclusion, design recommendations and further research activities are derived.

2. Materials & Methods

The TL and CL approaches are strategies of machine learning (ML), which in turn is a subdomain of artificial intelligence. The aim of ML is to enable machines to recognize patterns and develop appropriate solutions [6]. In the field of image recognition, especially convolutional neural networks (CNN) are employed. With these, for instance, defects in textiles can be detected, yet a lot of image data is required. Currently, simple data augmentation operations (e.g., flipping or rotating) are commonly pursued to artificially increase the amount of data. Complimentary to data augmentation, the intention of TL and CL is to utilize the existing data more efficiently and thus enable the CNN to achieve faster learning success [7]. The dataset used and the two ML approaches are explained in more detail below.

2.1 Dataset

The used dataset includes 3,653 image captures of biaxial ±45° glass non-crimp fabrics (fringe) with 320 g/m² grammage. The captured images have a resolution of 2,048 x 1,536 pixels and were acquired during three-dimensional forming of the textiles with an Apodius HP-C-V3D vision sensor, mounted on a Hexagon ROMER Absolute Arm. The forming was performed on the shape shown in Figure 1 as it favours multiple defects due to its complex corners and varying curvatures.

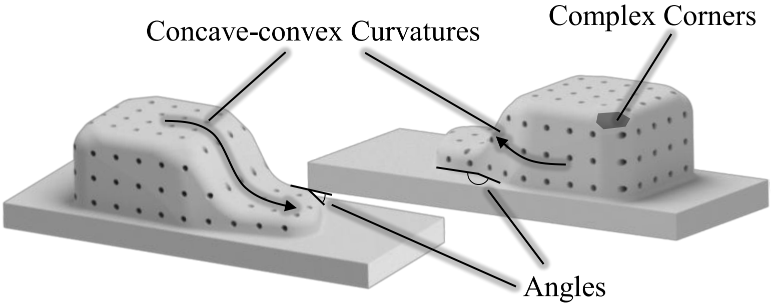


Figure 1: Geometry used for dataset generation

The raw images were segmented into 12 overlapping images with a resolution of 512 x 512 pixels each (Figure 2) and were then assigned to exactly one of the defect classes fold, flawless, gap, undulation, sling and distortion. Exemplary embodiments of each class are depicted in Figure 3. As some classes occur more frequently than others, data augmentation is used to create a balanced dataset. An unbalanced dataset could result in learning biased correlations and thus leading to erroneous classifications. Data augmentation is performed using the operations rescale, flip, brightness adjustment, rotation and zooming to create three distinct datasets with 500, 2,000 and 3,000 images for each class. All three datasets are used during a hyperparameter optimisation to determine a fitting network architecture as well as the best performing dataset. Hyperparameters are parameters that are not influenced by the dataset during training and therefore have to be specified before the network's training. The network's architecture is primarily determined by these hyperparameters, which are iteratively identified in a so-called hyperparameter optimisation. During our hyperparameter optimisation (HPO) we alter epochs, dropouts, learning rate, number of layers, number of neurons and batch size. After HPO, a validation dataset unknown to the network is applied to evaluate the generalisation capability (validation accuracy) of the network. The best performing augmented dataset is used for transfer and curriculum learning.

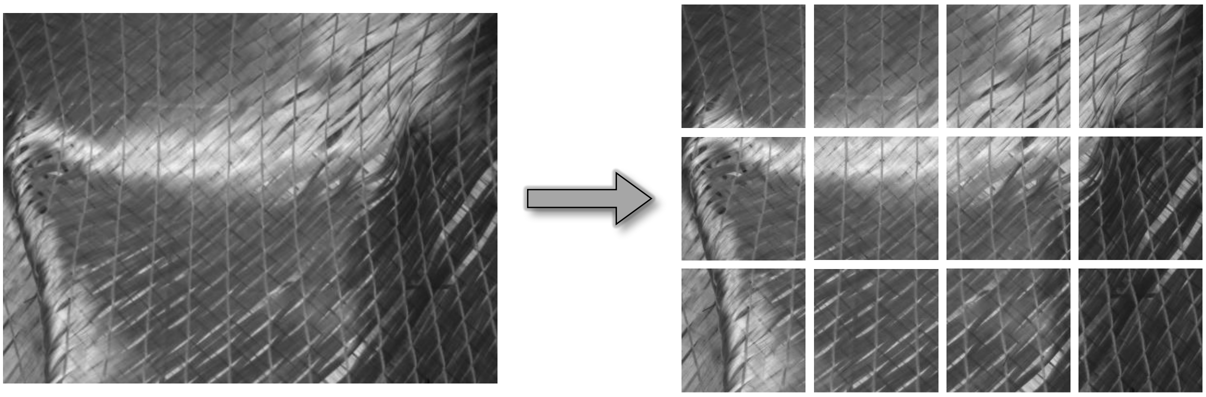


Figure 2: Segmentation of raw images into 12 parts

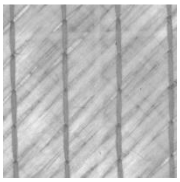
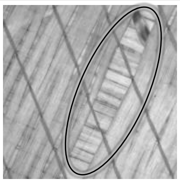
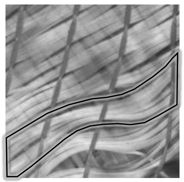
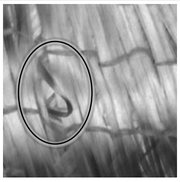
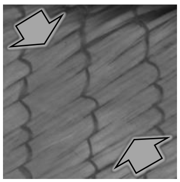
Fold	Flawless	Gap	Undula-tion	Sling	Distortion
					

Figure 3: Examples for each defect class

2.2 Transfer Learning

In transfer learning, the applied neural network is not designed from scratch, but a pre-trained network is used as a starting point. In the pre-trained network, the weights between the individual neurons are already preset. Starting from this, the weights can be adjusted more quickly and thus high validation accuracies can be achieved in fewer epochs [8].

In this contribution, the Xception [9] and VGG16 [10] networks are used for transfer learning because they exhibit high top 5 accuracy while having widely varying depth, data size, and parameter specification. A version of each network pretrained with the ImageNet dataset is used, a hyperparameter optimization is performed and the existing network architecture is specifically adapted. Based on this, design guidelines for transfer learning are derived and the validation results are compared with those of the regular (vanilla) network.

2.3 Curriculum Learning

Curriculum learning is a learning concept for neural networks in which content is processed in increasing complexity analogous to human teaching. In the context of neural networks, this means that during training not all data is fed to the network at once, but that the training dataset is gradually extended in increasing complexity until all data is included. Bengio et al. were able to demonstrate a performance improvement of a network by applying curriculum learning in different use cases [11]. However, this has not yet been done in the context of defect detection within the fibre-reinforced composite domain.

For CL, the regular trained vanilla network is used. Based on this, three different curricula are investigated in three variants each and then an HPO is performed for each curriculum. In Curriculum 1, the number of classes to be learned is gradually increased from 3 to 6 in four stages during training (Table 1). In Curriculum 2, six classes are introduced and differentiated from the beginning. Initially, this curriculum uses only images with horizontal fibres. In the course of the curriculum the diversity of the fibre trajectories increases. In addition, the defect characteristics of folds and loops are subjectively divided into four difficulty levels. A third variation of Curriculum 2 has an increased amount of steps (eight) in which the data is fed gradually during the network's training. The third curriculum is partitioned based on the illumination intensities of the images. In the first stage, bright images are fed, in the second stage dark images are added and in the third stage images with bright and dark areas are additionally provided, which have so-called shadow edges. In the last stage, the remaining images are fed. The three variants of each curriculum change the order in which classes, illumination intensities, or fibre orientations are fed to the training. These are randomized as well as descending and ascending according to their brightness (Curriculum 3), fibre orientation (Curriculum 2) or precision of each class of the vanilla mesh (Curriculum 1). Precision of a class is defined as the number of correctly labelled images of a class divided by the sum of correctly labelled images and the falsely allocated images to this class.

Table 1: Overview of examined curriculum strategies

	First stage of training	**Changes for following stages**
Curriculum 1	Starting with three classes	Adding one class every stage
Curriculum 2	Six classes, starting with horizonzal fibres only, subjective difficulty for loops and folds	Increasing deviation of fibre trajectories, increasing difficulty for loops and folds
Curriculum 3	Only bright images are used	Decreasing illumination within pictures

3. Results

Following, the results of the vanilla network's as well as the transfer and curriculum learning networks' HPO and validation are described.

3.1 Vanilla Approach

In the course of the vanilla approach, a CNN architecture is determined and all data is fed to the vanilla network at once during training. The HPO provides the highest validation accuracy and lowest loss function with the dataset consisting of 3,000 images. This dataset is subsequently used for curriculum and transfer learning. The HPO of the vanilla network results in a validation accuracy of 90.3 % and a loss value of 0.314.

The network consists of five convolution and one flattening layer, each followed by a dropout layer to avoid overfitting. The hyperparameters are 64-128-128-256-512 neurons per layer, 25 training epochs, a dropout value of 0.2, a learning rate of 0.0002 and a batch size of 20. The vanilla network's validation accuracy of 90.3 % is used as reference value.

3.2 Transfer Learning Approach

An HPO is performed with the Xception ImageNet. Analogous to the vanilla training, epochs, batch size, dropout, learning rate, number of layers and number of neurons per layer are varied. None of the 142 trained networks shows comparable convergence or accuracy as the vanilla network. Either the training and validation accuracies are lower than 75 % or a difference between validation and training functions is evident (overfitting). Since none of the trained networks has a sufficiently high accuracy, the approach with the Xception network is not investigated further.

Analogously, an HPO with the same parameters is also performed with the VGG16-ImageNet [10]. The highest validation accuracy of 93.06% with a loss of 0.5147 provides 15 epochs, a batch size of 5, a 0.3 dropout and a learning rate of 0.0002 (Figure 4). The number of layers and neurons per layer from [10] can be confirmed during the HPO, resulting in only adding one dropout layer at the end of the model and adjusting the output layer accordingly.

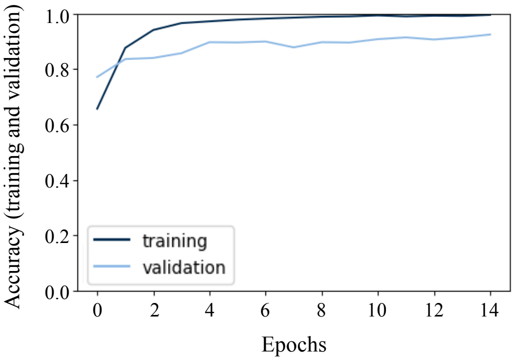


Figure 4: Training and Validation for VGG 16's Transfer Learning

Since the VGG16 is a very deep network with 13 convolutional layers, we omitted and added individual layers to investigate how this affected training. In total, we removed 2, 4, 6, and 9 layers and trained the network. For all networks with a reduced number of layers, a significant difference between training and validation data is evident, suggesting overfitting. The more layers removed, the larger the difference. In contrast, adding 2, 4, 6, and 9 layers shows an increase in the generalisation ability of the network. The difference between validation and training data decreases as the number of layers increases. From 6 additional layers onwards, the effect reverses and a reduction in validation accuracies and an increase in loss occurs.

3.3 Curriculum Learning Approach

The HPO results in the same hyperparameters being applied for Curriculum 1 as in the vanilla network. The three variants achieve a validation accuracy of 87.15 %, 91.07 % and 95.25 %. The highest validation accuracy (95.25 %) with a loss of 0.2202 is achieved by the curriculum in which the classes are gradually added to the training with increasing precision. Compared to the vanilla approach, the curriculum increases the precision of the class flawless from 74.17 % to 96.14 % during training and 90.39 % during validation. All other classes also show values above 90 %. In total, fewer images are misclassified. (Figure 5).

Predicted Class
(Vanilla)

Actual Class

	FO	**FL**	**GA**	**UN**	**SL**	**DI**
FO	2,113	7	2	0	0	25
FL	1	2,036	0	0	0	8
GA	21	359	1,680	1	3	44
UN	1	234	3	1,709	4	188
SL	27	55	0	0	1,878	134
DI	3	54	0	0	0	2,082

Predicted Class
(C1-V3)

	FO	**FL**	**GA**	**UN**	**SL**	**DI**
FO	2,147	0	0	0	0	0
FL	0	2,045	0	0	0	0
GA	1	62	2,042	1	1	1
UN	0	18	0	2,129	0	1
SL	0	2	0	0	2,092	0
DI	0	0	0	0	0	2,139

FO – Fold, FL – Flawless, GA – Gap, UN – Undulation, SL – Sling, DI - Distortion

Figure 5: Confusion Matrices during Training of Vanilla Approach (Vanilla, left) and Curriculum 1 Variant 3 (C1-V3, right)

For Curriculum 2, which is structured according to fibre orientations and defect characteristics, 25 training epochs, a learning rate of 0.0002, a dropout of 0.2 and a batch size of 20 are derived from HPO. With these parameters, validation accuracies of 95.00 % (random order), 93.06 % (decreasing fibre orientation and precision) and 90.79 % (increasing fibre orientation and precision, 8 stages) are achieved (Figure 6). The class flawless together with the class distortion shows the most classification errors and thus the lowest precision. However, the curriculum can increase the precisions to 89.46 % (distortion) and 93.17 % (flawless). Especially in the first stages of the curriculum, many images are misclassified as flawless or distortion during training, while other classes show less or hardly any classification errors. Accordingly, the precision values of flawless and distortion are below 60 % during the early stages.

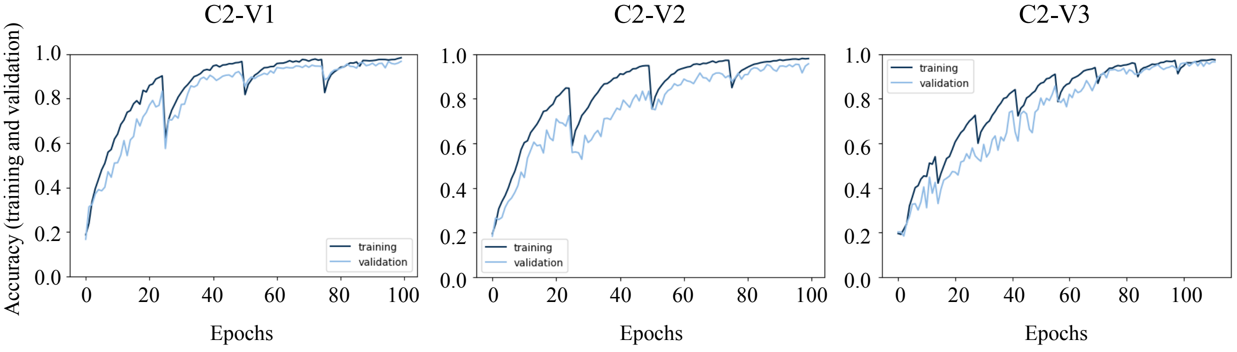


Figure 6: Accuracies of Training and Validation of Curriculum 2's (C2) three variants (V1-V3, left to right)

Analogous to the other two curricula, an HPO is also performed for the brightness-based Curriculum 3. The resulting hyperparameters are 50 training epochs, a learning rate of 0.0002, a dropout of 0.2 and a batch size of 20. All three variants achieve comparable validation accuracies, differences are in the per mil range (V1: 94.59 %, V2: 94.71 %, V3: 95.34 %).

Compared to the other two curricula, Curriculum 3 achieves the highest validation accuracy overall, but the precision of the class flawless with 84.38 % is lower than for Curriculum 1 (90.39 %) and Curriculum 2 (93.17 %). Consequently, the classification errors in Curriculum 1 and Curriculum 2 are distributed more evenly across all classes, while in Curriculum 3 the class flawless is the most frequently specified class for classification errors. Overall, higher validation accuracies are achieved with all three curricula ($Acc_{Vanilla}$ = 90.3 %; $Acc_{Curricula}$ > 94.59 %). The precision values for the frequently incorrectly specified class flawless are increased as well due to the curricula ($Prec_{Vanilla}$ = 74.17 %; $Prec_{Curricula}$ > 84.38 %).

4. Discussion & Conclusion

The results of transfer and curriculum learning indicate that both strategies can improve the generalisation ability of a convolutional neural network. Overall, fewer classification errors occur when one of the two approaches is employed. However, especially the defect class flawless seems to be difficult to distinguish from the other classes and contains the most incorrect classifications. In our opinion, the difficulty of differentiation is due to the great similarity to images of the class gaps. Images of the class flawless often have small spaces that can be misinterpreted as gaps. We assume that this is why gaps are most often incorrectly categorised as flawless. Interestingly, flawless images are not or hardly ever assigned to the class gaps. A similar observation can be made for the class distortion, where we attribute the incorrect classifications to superpositions of several defects that may additionally arise due to the distortion, thus posing further challenges to the network. Nevertheless, due to transfer and curriculum learning, the networks exhibit accuracies, precision and recalls of over 90 %.

In the context of the TL, pretrained networks (with ImageNet) can in some cases be adopted one-to-one for the reinforcement textile defects dataset. Whether this is possible depends strongly on the use case and the respective network structure. The more convolutional layers a pre-trained network has, the more data is needed to adapt all weights to the new use case in the course of training. Therefore, if the goal is to achieve maximum training success of the network with as little data as possible, lean networks with few convolutional layers should be selected. However, the results of the TL suggest that additional layers enable better feature recognition and thus contribute to higher classification accuracies. At this point, further research is needed to investigate causal relationships between network depth, further hyperparameters and validation accuracies. Based on these yet to be gained insights, elaborated design recommendations for TL in production related context can be derived in the future.

In curriculum learning, the results show that the derivation of a semantically designed learning plan has a beneficial effect on the network's performance indicators (accuracies, precision, recall). However, it is not only the semantic, stage-by-stage division into a curriculum that is important, but also the order in which the stages are presented to the network. The analysis of the three curricula shows that, for example, the arrangement according to precision or the arrangement according to the subjective perception of difficulty apparently has no immediate correlation on an improvement. These approaches are therefore not suitable as general design recommendations for a curriculum. In this context, the ethically and morally motivated question arises whether a machine learns in the same way as a human being and is able to perform a nuanced differentiation? Based on this question, we propose to develop metrics for the perception of difficulty or relatedness of data units. We see initial starting points for this in the use of confidence scores as well as density-based clustering approaches to measure difficulty and identify semantically related data points.

In conclusion, we observe that both learning strategies achieve higher accuracies in training and validation as well as higher discriminatory power in classification. The observed learning effect is comparable to an increase in the amount of data in the vanilla approach. As a result, both CL and TL contribute to making big data approaches accessible for applications with few data available. Thus, the further investigation and successful enhancement of both approaches represents an essential milestone in making artificial intelligence accessible in data scarce environments.

Acknowledgements

Funded by the Federal Ministry of Education and Research (BMBF) in context of the project “KIQS - AI based quality control system to classify draping defects” (01IS20089C).

References

[1] Towsyfyan, H., Biguri, A., Boardman, R., Blumensath, T., 2020. Successes and challenges in non-destructive testing of aircraft composite structures. *Chinese Journal of Aeronautics*, *33*(3), 771-791. doi: 10.1016/j.cja.2019.09.017

[2] Fuhr, J. P., 2017. *Schichtbasierte Modellierung von Fertigungseffekten in der Struktursimulation von Faserverbundwerkstoffen*. Verlag Dr. Hut. Dissertation

[3] Luczak, H., Mueller, T., 1994. Worker attention and fatigue. *Design of work and development of personnel in advanced manufacturing*, 463-492.

[4] Tekin E., Kapan Ö., 2016. Composite manufacturing data management in aerospace industry. *Procedia CIRP 41*, pp. 1039-1042, doi: 10.1016/j.procir.2015.12.058

[5] Perez, L., Wang, J., 2017. The effectiveness of data augmentation in image classification using deep learning. *arXiv preprint arXiv:1712.04621.*

[6] Erickson, B. J., Korfiatis, P., Akkus, Z., Kline, T. L., 2017. Machine learning for medical imaging. *Radiographics*, *37*(2), 505-515. doi: 10.1148/rg.2017160130

[7] Shorten, C., Khoshgoftaar, T. M., 2019. A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 1-48. doi: 10.1186/s40537-019-0197-0

[8] Neyshabur, B., Sedghi, H.,, Zhang, C., 2020. What is being transferred in transfer learning?. *arXiv preprint arXiv:2008.11687.*

[9] Chollet, F., 2016. Xception: Deep Learning with Depthwise Separable Convolutions. *arXiv preprint arXiv:1610.02357.*

[10] Simonyan, K.; Zisserman, A., 2014. Very deep convolutional networsk for large-scale image recognition. *arXiv preprint arXiv:1409.1556.*

[11] Bengio, Y., Louradour, J., Collobert, R., Weston, J., 2009. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning* (pp. 41-48). doi: 10.1145/1553374.1553380

Biography

Florian Brillowski is an engineer specialized in plastics processing and operations research. He is working at the Institute for Textile Technology of RWTH Aachen University as research associate since 2018. His research interests are AI, digitisation, automation, composites and prototyping. He is currently working on his dissertation about the acceptance of AI-based decision support systems.

Vanessa Overhage is a textile engineer specialised in composite materials. Since 2021 she is a research associate at the Institute of Textile Technology at RWTH Aachen University. Her research interests are composites, sustainability and AI. She is currently esearching the use of recycled carbon fibres in the construction industry.

Thorsten Tegetmeyer-Kleine graduated in electrical engineering from RWTH Aachen University with major in systems and automation. During his master's degree in 2019, he started specializing in Deep Learning. Currently, he is employed as Software Developer for AI at MaBri Vision GmbH

Benjamin Hohnhäuser is the head of the robotics research group, department Image Processing / Industrial Applications at the Society for advancement of applied computer science (GFaI e.V.). His major fields of research are 3D-data processing, automated coating, machine learning and robotics.

3rd Conference on Production Systems and Logistics

Computer-Aided Assembly Sequence Planning For High-Mix Low-Volume Products In The Electronic Appliances Industry

Mingyue Yao[1], Benny Drescher[1,2], Erik Stewart [1], Lakshmi Manoj [1], Max Wittstamm [1,3], Leonhard Henke [4], Shyam Ghodasara [5], Ming Ge[1]

[1]Hong Kong Industrial Artificial Intelligence and Robotics Centre (FLAIR), Hong Kong, SAR China
[2]Invention Center at RWTH Aachen Campus (INC), Aachen, Germany
[3]Center Connected Industry at RWTH Aachen Campus (CCI), Aachen, Germany
[4]FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany
[5]Center Integrated Business Applications at RWTH Aachen Campus (CIBA), Aachen, Germany

Abstract

Electronic appliance manufacturers are facing the challenge of frequent product orders. Based on each product order, the assembly process and workstations need to be planned. An essential part of the assembly planning is defining the assembly sequence, considering the mechanical product's design, and handling of the product's components. The assembly sequence determines the order of processes for each workstation, the overall layout, and thereby time and cost. Currently, the assembly sequence is decided by industrial engineers through a manual approach that is time-consuming, complex, and requires technical expertise. To reduce the industrial engineers' manual effort, a Computer-Aided Assembly Sequence Planning (CAASP) system is proposed in this paper. It compromises the components for a comprehensive system that aims to be applied practically. The system uses Computer-Aided Design (CAD) files to derive Liaison and Interference Matrices that represent a mathematical relationship between parts. Subsequently, an adapted Ant Colony Optimization Algorithm generates an optimized assembly sequence based on these relationships. Through a web browser-based application, the user can upload files and interact with the system. The system is conceptualized and validated using the CAD file of an electric motor example product. The results are discussed, and future work is outlined.

Keywords

Assembly sequence; production planning; computer-aided assembly planning; industry 4.0; High-Mix Low-Volume

1. Introduction

Assembly planning describes the planning of bringing individually machined parts into a final product of higher complexity [1]. It incorporates the planning of the assembly tasks and required tools, the sequence, the layout, and resources [2]. Due to the market demands for product versions and models, smaller batch runs, and shorter concept-to-market lead times, the effort of assembly planning is rising [3–5]. The importance of assembly planning is high as assembly accounts for up to 70% of the production costs [1]. The assembly sequence planning (ASP) is a fundamental step of the assembly planning process. It has a significant impact on the assembly process, such as the assembly line layout and the operations at each workstation, and thereby production efficiency and cost [6]. It is complicated to find the optimal assembly sequence out of the vast number of possibilities [7]. ASP remains a manual task due to complexity reasons

DOI: https://doi.org/10.15488/12149

and the impact on the assembly process. The automatization of ASP has been examined in research as one of the main drivers for assembly optimization [8].

It can be observed that especially electronic appliances contract manufacturers with a high-mix low-volume production strategy can profit from an automated ASP solution because of the high need and feasibility. The need to automate ASP arises on the one side due to the increasing assembly planning effort. Electronics contract manufacturers apply the high-mix low-volume production method as a reaction to market demands, wide client base and a need to increase profitability [9]. High-mix low-volume is characterized by low unit volumes and high diversity, resulting in increased assembly planning effort [10]. Besides, contract manufacturers are under constant pressure to reduce costs [9]. Finally, the ASP for contract manufacturers is complicated by the limited connection between product design and assembly planning [11]. On the other side, the feasibility of an automated ASP is high. ASP can be automated more easily because of the increasing uniformity of products due to digitization, that help contract manufacturers serve various customers while still achieving economies of scale. This is especially relevant for electronic products, as the parts are primarily standard parts like resistors, capacitors, memory chips, which are assembled in many different product configurations [12,13]. Also, it can be derived that the high share of manual assembly operations supports the feasibility. Electronic appliance assembly mainly consists of the final product assembly which is dominated by manual operations [14]. The planning complexity and resulting costs of manual assembly are lower in contrast to automated assembly systems [15].

Whilst theoretical approaches have been outlined in scientific literature, practical industry solutions for an automated ASP remain limited [16,17]. For instance, common CAD software offer minimal support in determining the assembly sequence [18]. The paper describes an approach, named Computer-Aided Assembly Sequence Planning (CAASP), for high-mix low-volume products in the electrical appliances contract manufacturers industry. An overview of related work in assembly sequence planning is given (section 2). The solution approach is presented (section 3), and the solution is validated by an electric motor assembly example (section 4). Finally, a conclusion is derived, and future work is outlined (section 5).

2. Related work

The related work on ASP can be structured regarding practical and theoretical approaches [19], which are described in more detail in the following sections.

2.1 Practical approaches

There are different practical approaches for assembly sequence planning. Such technological solutions must be applicable in the industry and pursue the goal of facilitating assembly sequence planning for the industrial engineer. Existing practical approaches are explained below. Based on the necessary information for the assembly sequence planning, precedence rules and graphs can be generated [20,21]. These approaches are further developed, for example by *Hao et al.* [22] using a genetic algorithm combined with the simulated annealing algorithm to search for the disassembly sequence planning. However, these approaches have in common that they require the intervention of assembly planners to gather further information such as additional precedence relations of subassemblies. [19] Due to the increasing use of CAD systems, the extraction of information from CAD files to generate an assembly sequence emerged as a field of research. However, such approaches as those of *Mathew et al.* [23] are still characterized by manual efforts, e.g., in the form of manual labelling or quality problems of the results. To further reduce manual efforts *Hadj et al.* [24] developed an add-in for the CAD-software SolidWorks that is used in the design phase to increase the efficiency of product development processes by exclusively considering feasible assembly sequences. However, applicability in terms of consideration of other CAD software and convenience for users are not the focus of these approaches. *Gulivindala et al.* [25] concludes that the information distribution of assembly sequences and the practical feasibility is not given. Based on these findings, a cloud-based solution for automatic disassembly planning with a genetic algorithm is developed, which results can be efficiently distributed to Internet of Things devices. However, since this research does not focus on the assembly sequence problem during the product development phase and the corresponding use for assembly planners, the previously mentioned problems of using different CAD software and the usability for assembly planners remain. It appears that so far no sufficiently practicable approach has been found that meets today's requirements for user-centeredness and system independence. To achieve practicability, the CAASP is based on a system-independent architecture without installation effort and an intuitive user interface.

2.2 Theoretical approaches

Currently, methods for generating assembly sequences can be mainly divided into two categories – mathematically based and artificial intelligence (AI) based. Mathematically based methods use diagrams, graphs, or matrices to generate assembly sequences, while AI methods are used to generate optimal assembly sequences. On the one hand, precedence diagrams [26,27] and liaison diagrams [28,29] were originally used to describe part relationship in generating assembly sequences. However, those diagrams need to be generated manually. The manual work of creating the matrices was automatized. This was supported by use of CAD software, where part and assembly information is available in digitized data format. This provides foundation to automatically generate mathematical models and opens new era for assembly sequence planning. *Gu and Yan* present an approach that automatically disassembly sequences based on connectivity diagrams using CAD data from a feature-based data base [30]. *Hadj et al.* used mating data extraction and collision analysis to generate assembly sequences automatically, directly integrated in CAD software by using its application programming interface [19]. Although the manual effort could be reduced tremendously, mathematical methods can only generate feasible assembly sequences, it cannot generate optimal assembly sequences and lacks practical usage which hasn't been tested on complex products.

On the other hand, AI methods, e.g., genetic algorithm [22], neural networks [31], particle swarm optimization [32], artificial immune systems [33] have been used for automatic assembly sequence planning. *Lu* and *Yang* [34] used ant colony algorithm for ASP, but it needs human intervention as assembly task priority diagram needs to be generated manually. *Huang* and *Xu* [6] combined mathematical methods with AI methods to use integrated disassembly interference matrix, connection matrix, integrated support matrix and Ant Colony Optimization (ACO) to solve assembly sequence problems. However, the matrices are generated manually and require human intervention. *Pan et. al* proposed an automatic way for assembly sequence planning which firstly introduced input as STEP CAD files [35]. The method extracts geometrical information for interference-free matrices which represents interference relationships between assembly components and then automatically generate assembly sequences with minimum number of assembly direction reorientations. However, it has only been tested for products with less than 5 components and for complex products with large number of components, the performance cannot be guaranteed.

Although new ear of automatic assembly sequence planning method opens, fully automatic assembly sequence planning system with user interface which can be applied to complex products still needs to be developed. Therefore, in this paper, a fully automatic assembly sequence planning system which doesn't need human intervention together with user interface is proposed to show the advance. It combines and adapts existing methods. The automatic assembly sequence planning system shall reduce manual work in assembly sequence planning, and the generated assembly sequence shall provide vast support to the industrial engineers for assembly planning work which in the end can save time and improve efficiency in a production environment.

3. Computer aided-assembly sequence planning (CAASP) system

In the following, the CAASP system is described. It handles the flow of information starting from the user upload, see Figure 1. The relevant product data is extracted (phase 1), described in section 3.1. The geometrical constraints are modelled (phase 2) the optimal assembly sequence is generated (phase 3), described in section 3.2. Finally, the results are visualized to the user.

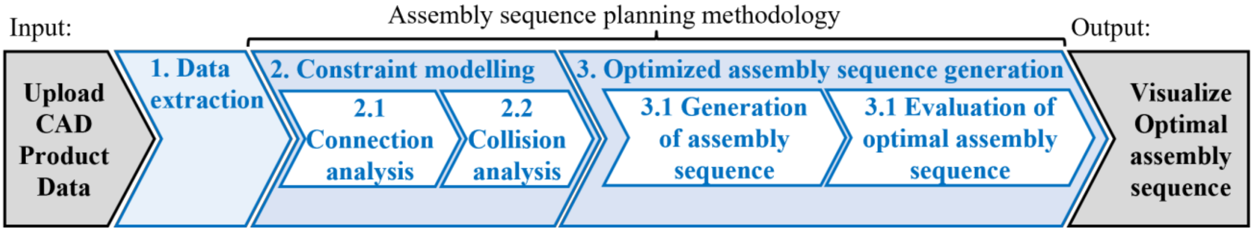


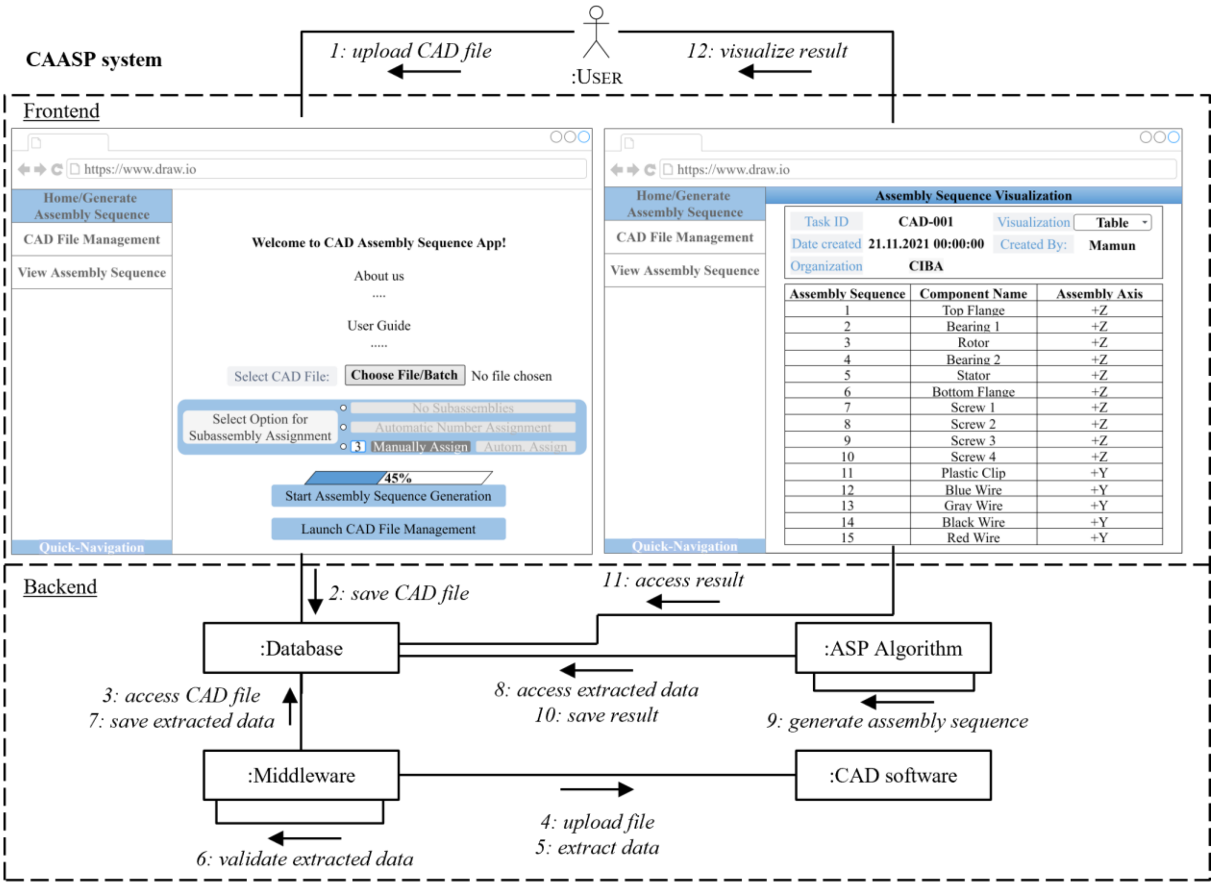
Figure 1: CAASP system flow of information

3.1 CAASP software system description

The ASP methodology is embedded in the ASP algorithm in the CAASP system to facilitate the usage by the user, e.g. an industrial engineer. The process is shown in a system diagram in Figure 2, modelled in the Unified Modelling Language (UML) [36]. The user uploads a CAD file in a web-based and user-centred application. This file is stored in a database with file management capabilities, along with additional information such as the date the file was uploaded. The database assigns a task ID to this file which is used as a unique identification number for the backend. The middleware is a backend component which communicates between database and CAD software. The middleware automatically extracts additional data necessary for assembly sequence planning from the CAD file by connecting to CAD software using application programming interfaces (APIs). Hence, manual effort to collect necessary data for assembly sequence planning can be significantly reduced. The data extraction was developed for the common STEP format, the Standard for the Exchange of Product model data [37]. The STEP format opens the opportunity to use various CAD software for the data extraction purpose and thus realize an agnostic and vendor independent approach. After extraction the data is validated and stored in the database. The ASP algorithm accesses this data and generates an assembly sequence, see section 3.2. The results are stored in the database so that the optimal assembly sequence results can be accessed by continuous pull requests from the frontend and displayed in a practical way for the user, referred to as the component name.

3.2 Assembly sequence planning methodology

The detailed methodology, consisting of two phase is presented in Figure 3. The first phase is the modelling of the geometric constraints as input for the optimization. The CAD input data is enriched to derive geometrical constraints like spatial data, part relationships, and collision information. In the CAASP system, the liaison matrix is applied to analyse the connections (step 2.1), while the interference matrices are utilized to analyse the collisions between the parts CAD file (step 2.2), see Figure 3. The liaison matrix represents contact information between two parts in an assembly and is produced by examining the connections between every part in a file [23]. The connected parts have a value of '1', while '0' indicates no connection between two parts. An interference matrix is also produced from the assembly file using collision analysis [19]. The assembly parts that interfere with other parts along the +x, -x, +y, -y, +z, and -z axes are identified. The information along the six axes is stored in six different binary matrices with values ‘0’ and ‘1’ where '0' indicates that there is no collision and ‘1’ indicates collision [38]. Step 3.1 follows, which is the generation of assembly sequences. These are produced by generating a disassembly sequence, which is then reversed to produce an optimal assembly sequence. For more complex assemblies with sub-assemblies, optimal sequences are generated for each of the subassemblies, after which a sequence is generated for the entire assembly.



Task ID CAD-001
Visualization Table
Date created 21.11.2021 00:00:00
Created By: Mamun
Organization CIBA

Assembly Sequence	Component Name	Assembly Axis
1	Top Flange	+Z
2	Bearing 1	+Z
3	Rotor	+Z
4	Bearing 2	+Z
5	Stator	+Z
6	Bottom Flange	+Z
7	Screw 1	+Z
8	Screw 2	+Z
9	Screw 3	+Z
10	Screw 4	+Z
11	Plastic Clip	+Y
12	Blue Wire	+Y
13	Gray Wire	+Y
14	Black Wire	+Y
15	Red Wire	+Y

Figure 2: Frontend and backend of CAASP system

The ASP problem is first converted into a graph with nodes and edges as mentioned in Figure 3, step 3.1. The nodes are created by combining a part number and potential orientations along the six axes. For example, part 1 would have the nodes (1, ‘+x’), (1, ‘-x’), (1, ‘+y’), (1, ‘-y’), (1, ‘+z’), and (1, ‘-z’) [39]. For the purpose of this research, minimal orientation change between parts has been identified as the primary requirement of an optimal assembly sequence. Each edge is initialized with a weight based on orientation change between two nodes that connect the edge. Angular changes of 0°, 90°, and 180° will results in a weight initialization of 1, 5, and 10 respectively. For example, an edge connecting (1, ‘+x’) and (4, ‘+x’) has a weight of 1, (1, ‘+x’) and (4, ‘-y’) has a weight of 5, and (1, ‘+x’) and (4, ‘-x’) has a weight of 10. Since the ASP problem can be represented as a graph with nodes and edges, the ACO method can be utilized to obtain a solution. Based on the logic of the algorithm, pheromone levels between all nodes are initialized with a concentration level The potential starting points for a sequence are identified by searching for rows with all zeroes in the interference matrix (step 3.1). All zeroes in a row of a part implies that the part is not blocked by other parts during disassembly. The number of ants is initialized to the number of starting points, and the ants are randomly placed at these points.

The following step is to identify the next feasible disassembly node for the ant. Out of the remaining components to be disassembled, parts that are not blocked by any of remaining components are selected using the interference matrix. The liaison matrix is then used to isolate parts from the selection that are in contact with the remaining components. If no parts are in contact, then the components chosen based on the interference matrix are directly utilized. Node selection from these potential nodes is done using the ACO probability formula as in [40]. The heuristic component in the algorithm becomes the weights, which are based on orientation initialized in the graph. Once the next part is selected based on the probability function, the algorithm checks if all the parts have been visited by each ant. If there are parts that still need to be visited by each ant, the node selection process is repeated. When all the parts have been visited, the pheromone levels for each path is incremented using the pheromone updating formula [40]. Once all the ants have completed an iteration, the best sequence of the iteration is selected based on the least number of reorientations (step 3.2). Subsequently, the pheromone evaporation is performed on all the edges using the evaporation formula in [40]. If iterations are still to be completed, the entire process is repeated. The optimal sequence is selected based on the global pheromone levels after all the iterations have been completed.

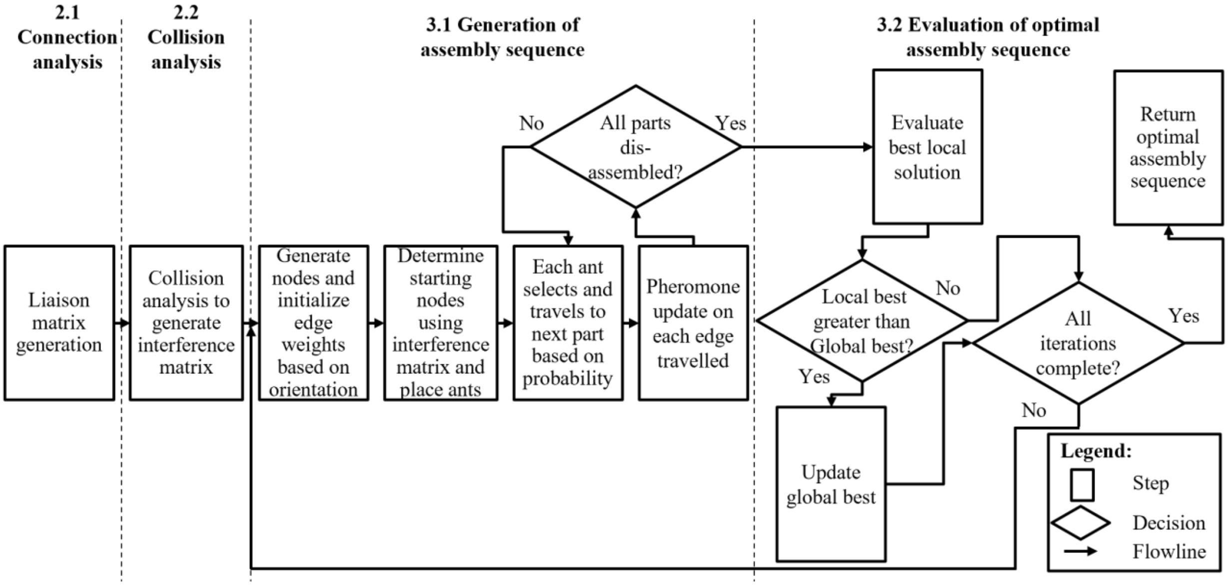


Figure 3: Detailed view of ASP methodology

4. Validation

The CAASP system is currently set up to automate the data processing between the user upload and the presentation of the result. To verify the practicability and accuracy of the methodology discussed, a 3D STEP file motor consisting of 15 components was chosen for validation [41] of the ASP methodology described in section 3.2. Figure 4 shows the CAD model, which represents a typical product in the Electronics Appliance Industry produced by contract manufacturers. The model contains 15 components that each have various contacts in the liaison matrix and multiple collisions along axes in the interference matrices. This model should ideally be assembled by dividing its components into two groups that are assembled along two different axes. Therefore, this motor model provides a relevant opportunity for the ASP methodology to optimize the assembly sequence by considering the minimization of component orientation changes during the assembly process.

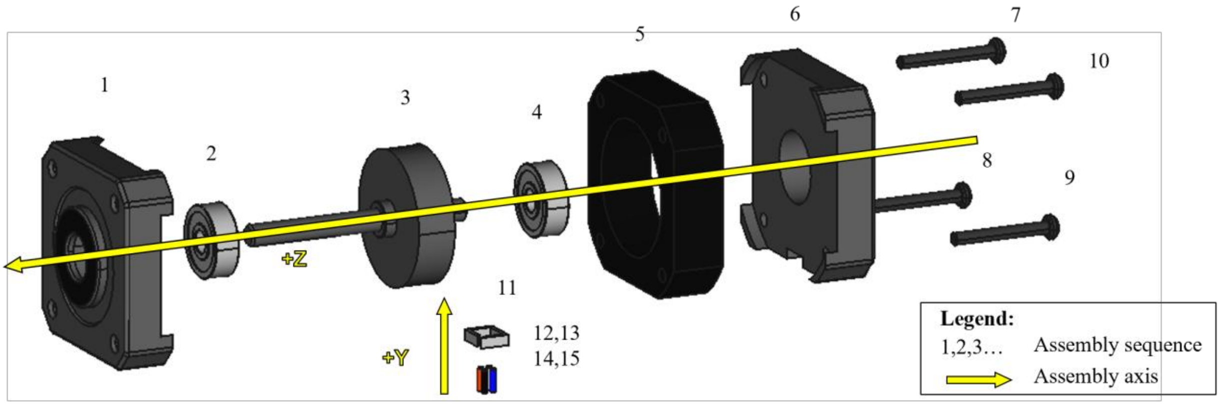


Figure 4: Labelled assembly sequence with directional axes (example)

The integrated system has been utilized to process a STEP file as input for the analysis. After the file is read, the system automatically calculates the liaison matrix, six collision matrices, and generates an optimal assembly sequence. The liaison matrix and collision matrix results are shown in Figure 5. Although there is one matrix for each of the 6 axes produced in the collision analysis results, only the +Z axis matrix is shown below for reference. The values in the first row and column ranging from 1 to 15 represent the components of the motor assembly, which can be referenced in the "Component Number" row of Table 1. The Component Number is used in this section for convenience of associating the values in the matrices with component numbers in the generated assembly sequence.

Liaison Matrix:

$$\begin{array}{c|ccccccccccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
7 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
8 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
9 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
10 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
11 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
15 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Interference Matrix: +Z Axis

$$\begin{array}{c|ccccccccccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
6 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
10 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\
11 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
12 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
14 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
15 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{array}$$

Figure 5: Liaison matrix and one interference matrix in the +Z axis

It should be mentioned that although every assembly sequence generated by CAASP system is mechanically feasible, a minor number of generated sequences is not practical since bolts or screws are chosen as the first component of the sequence. The focus of the following is on most sequences that are both feasible and practical. A noteworthy aspect to discuss is how the CAASP system created one group of components to be sequentially assembled in the +Z direction and one group of components to be sequentially assembled in the +Y direction, as shown in Table 1. Figure 4 provides a visual representation for the labelled assembly sequence of each component along the assembly axes (example result). The results shown represent one variation out of many similar, potential optimal assembly sequences. Since the ASP algorithm attempts to optimize the assembly sequence by minimizing assembly axis orientation changes, the assembly sequence first consists of components from the +Z group, followed by components from the +Y group, thus resulting in an assembly sequence with minimum orientation changes. Based on the optimization criteria, when comparing the algorithm-generated assembly sequence to the known optimum assembly sequence, it can be shown that both sequences achieve the same level of optimization from a feasibility and assembly orientation perspective.

Table 1: Assembly Sequence Result (example)

Category	Output														
Assembly Sequence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Component Number	9	8	4	6	7	10	5	13	15	3	12	1	2	14	11
Assembly Axis	+Z	+Z	+Z	+Z	+Z	+Z	+Z	+Z	+Z	+Z	+Y	+Y	+Y	+Y	+Y

The total runtime for the entire calculation takes 5.3 minutes on average. It should be noted that the liaison matrix calculation and optimized assembly sequence planning only require 15 seconds in total to complete, while the remaining 95% of runtime is spent on collision analysis for calculation of the liaison matrices. This highlights the need for further optimizations in the collision analysis method in future works.

5. Discussion and future work

The research work outlines a feasible solution for contract manufacturers to support the industrial engineers' planning of assembly sequences. The solution fulfils user-centeredness by a web-based frontend. System dependencies are reduced by a backend architecture that can be deployed company-internally or cloud-based. CAASP supports STEP format to build a vendor-agnostic solution. The methodology automates the ASP of contract manufacturers with minimal human intervention. The validation of the methodology shows the feasibility of the approach exemplified in the ASP of an electronic motor. Presently, the limitations are seen in the execution of CAASP as a high amount of computation time is needed for analysing part collisions. In the electronic motor example, this leads to a calculation time of 5.3 minutes for 15 parts in the liaison matrices. The industrial applicability of the approach is restrained by using the minimal orientation change between parts as a solemn optimization criterion. In the electronic motor assembly, this can result in proposing feasible but impractical sequences (minor cases). For example, the plastic clip precedes the wire assembly – although feasible, it is not practicable in an actual industrial setting. Future research aims at increasing the practicability of CAASP. It shall handle various optimization criteria and algorithmic constraints common in the electronic appliance industry. The CAASP system will be applied to several product types, models, and variants to adjust the system for wide usage in the industry. Furthermore, application engineering shall be conducted for seamless integration into business processes and high compatibility with state-of-the-art software vendors.

Acknowledgements

Our research is funded by AIR@InnoHK research cluster of the Innovation and Technology Commission (ITC) of the HKSAR Government. The results presented in this paper within the scope of the research project 'Digital Shadow for Self-Optimizing Electronics Assembly' has come to a successful completion because of the support from ITC. We would like to express our sincere gratitude to them.

The authors want to thank Dr. Ziyan Lu, Frankie Hui Ka Yiu, Katharina Berwing, and Md Al Mamun for their meaningful contribution to the concept and implementation of the solution.

References

[1] Lotter, B., 2012. Einführung, in: Lotter, B. (Ed.), Montage in der Industriellen Produktion. Ein Handbuch Für Die Praxis, 2. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg, pp. 1–9.

[2] Wildemann, H., 2010. Neue Montagekonzepte: Realisierung von Produktordnungssystemen in der Kleinserienmontage komplexer Produkte bei kleinen und mittleren Unternehmen. Research Report, München.

[3] Kern, W., Rusitschka, F., Bauernhansl, T., 2016. Planning of Workstations in a Modular Automotive Assembly System. Procedia CIRP 57, 327–332.

[4] Schuh, G., Prote, J.-P., Dany, S., Molitor, M., Giner, B., 2018. Process model for generative assembly planning in the highly iterative product development. Procedia CIRP 72, 363–368.

[5] Wiendahl, H.-P., El Maraghy, H.A., Nyhuis, P., Zäh, M.F., Wiendahl, H.-H., Duffie, N., Brieke, M., 2007. Changeable Manufacturing - Classification, Design and Operation. CIRP Annals 56 (2), 783–809.

[6] Huang, W., Xu, Q., 2017. Automatic generation and optimization of stable assembly sequence based on ACO algorithm, in: 2017 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 2057–2062.

[7] Bahubalendruni, M.R., Biswal, B.B., 2016. A review on assembly sequence generation and its automation. Journal of Mechanical Engineering Science 230 (5), 824–838.

[8] Deepak, B., Bala Murali, G., Bahubalendruni, M.R., Biswal, B.B., 2019. Assembly sequence planning using soft computing methods: A review. Journal of Process Mechanical Engineering 233 (3), 653–683.

[9] Lüthje, B., Butollo, F., 2017. Why the Foxconn Model Does Not Die: Production Networks and Labour Relations in the IT Industry in South China. Globalizations 14 (2), 216–231.

[10] Abele, E., Meyer, T., Näher, U., Strube, G., Sykes, R. (Eds.), 2008. Global Production: A Handbook for Strategy and Implementation. Springer Berlin Heidelberg, Berlin, Heidelberg.

[11] Tsai-Lin, T.-F., Chi, H.-R., Chang, Y.-C., 2021. The business model and innovation mix in the transition of contract manufacturers in the greater China region. Asia Pacific Business Review 27 (3), 444–469.

[12] Barnes, E., Dai, J., Deng, S., Down, D., Goh, M., Lau, H.C., Sharafali, M. Electronics Manufacturing Service Industry. Research Report. School of Industrial and Systems Engineering Georgia Institute of Technology.

[13] Morschett, D., Donath, A., 2011. Die Globale Strategie von Flextronics – Contract Manufacturer als Outsourcing-Partner international tätiger Industrieunternehmen, in: Swoboda, B., Morschett, D., Zentes, J. (Eds.), Fallstudien zum Internationalen Management, vol. 26, 4. ed. Gabler Verlag, Wiesbaden, pp. 211–228.

[14] Koh, D., Chan, G., Yap, E., 2004. World at work: the electronics industry. Occupational and environmental medicine 61 (2), 180–183.

[15] Hesse, S., 2012. Automatische Montagemaschinen, in: Lotter, B. (Ed.), Montage in der Industriellen Produktion. Ein Handbuch Für Die Praxis, 2. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg, pp. 195–272.

[16] Hadj, R.B., Belhadj, I., Trigui, M., Aifaoui, N., 2018. Assembly sequences plan generation using features simplification. Advances in Engineering Software 119, 1–11.

[17] Trigui, M., Ben Hadj, R., Aifaoui, N., 2015. An interoperability CAD assembly sequence plan approach. The International Journal of Advanced Manufacturing Technology 79 (9-12), 1465–1476.

[18] Gors, D., Put, J., Vanherle, B., Witters, M., Luyten, K., 2021. Semi-automatic extraction of digital work instructions from CAD models. Procedia CIRP 97 (11), 39–44.

[19] Hadj, R.B., Trigui, M., Aifaoui, N., 2015. Toward an integrated CAD assembly sequence planning solution. Journal of Mechanical Engineering Science 229 (16), 2987–3001.

[20] Dini, G., Santochi, M., 1992. Automated Sequencing and Subassembly Detection in Assembly Planning. CIRP Annals 41 (1), 1–4.

[21] Homem de Mello, L.S., Sanderson, A.C., 1991. A correct and complete algorithm for the generation of mechanical assembly sequences. IEEE Trans. Robot. Automat. 7 (2), 228–240.

[22] Hao, W., Hongfu, Z., 2009. Using genetic annealing simulated annealing algorithm to solve disassembly sequence planning. Journal of Systems Engineering and Electronics 20 (4), 906–912.

[23] Mathew, A.T., Rao, C.S.P., 2010. A Novel Method of Using API to Generate Liaison Relationships from an Assembly. JSEA 03 (02), 167–175.

[24] Hadj, R.B., Belhadj, I., Gouta, C., Trigui, M., Aifaoui, N., Hammadi, M., 2018. An interoperability process between CAD system and CAE applications based on CAD data. International Journal on Interactive Design and Manufacturing 12 (3), 1039–1058.

[25] Kumar Gulivindala, A., V. A. Raju Bahubalendruni, M., Chandrasekar, R., Ahmed, E., Haider Abidi, M., Al-Ahmari, A., 2021. Automated Disassembly Sequence Prediction for Industry 4.0 Using Enhanced Genetic Algorithm. Computers, Materials & Continua 69 (2), 2531–2548.

[26] Baldwin, D.F., Abell, T.E., Lui, M.-C., Fazio, T.L. de, Whitney, D.E., 1991. An integrated computer aid for generating and evaluating assembly sequences for mechanical products. IEEE Trans. Robot. Automat. 7 (1).

[27] Fazio, T. de, Whitney, D., 1987. Simplified generation of all mechanical assembly sequences. IEEE J. Robot. Automat. 3 (6), 640–658.

[28] Bullinger, H.J., Ammer, E.D., 1984. Computer-aided depicting of precedence diagrams—A step towards efficient planning in assembly. Computers & industrial engineering 8 (3-4), 165–169.

[29] Prenting, T.O., Battaglin, R.M., 1964. The precedence diagram: A tool for analysis in assembly line balancing. Journal of Industrial Engineering 15 (4), 208–213.

[30] GU, P., YAN, X., 1995. CAD-directed automatic assembly sequence planning. International Journal of Production Research 33 (11), 3069–3100.

[31] Suszyński, M., Peta, K., 2021. Assembly Sequence Planning Using Artificial Neural Networks for Mechanical Parts Based on Selected Criteria. Applied Sciences 11 (21), 10414.

[32] Mukred, J.A., Muslim, M.T., Selamat, H., 2013. Optimizing Assembly Sequence Time Using Particle Swarm Optimization (PSO). AMM 315, 88–92.
[33] Chang, C.-C., Tseng, H.-E., Meng, L.-P., 2009. Artificial immune systems for assembly sequence planning exploration. Engineering Applications of Artificial Intelligence 22 (8), 1218–1232.
[34] Lu, C., Yang, Z., 2016. Integrated assembly sequence planning and assembly line balancing with ant colony optimization approach. The International Journal of Advanced Manufacturing Technology 83 (1-4), 243–256.
[35] Pan, C., Smith, S.S., Smith, G.C., 2006. Automatic assembly sequence planning from STEP CAD files. International Journal of Computer Integrated Manufacturing 19 (8), 775–783.
[36] Object Management Group (OMG), 2017. OMG® Unified Modeling Language® (OMG UML®), 2nd ed. https://www.omg.org/spec/UML/2.5.1/PDF. Accessed 20.01.22.
[37] International Organization for Standardization (ISO), 2020. ISO 10303-242:2020: Industrial automation systems and integration - Product data representation and exchange - Part 242: Application protocol: Managed model-based 3D engineering, 242nd ed. 25.040.40. https://www.iso.org/standard/66654.html. Accessed 20.01.22.
[38] Bahubalendruni, M.R., 2016. Computer Aided Optimal Robotic Assembly Sequence Generation. Dissertation, Rourkela.
[39] Han, Z., Wang, Y., de Tian, 2021. Ant colony optimization for assembly sequence planning based on parameters optimization. Frontiers of Mechanical Engineering 16 (2), 393–409.
[40] Yang, J., Shi, X., Marchese, M., Liang, Y., 2008. An ant colony optimization method for generalized TSP problem. Progress in Natural Science 18 (11), 1417–1422.
[41] Ondřej Hynek, 2021. Microcon SX17-1003LQCEF. GRABCAD. https://grabcad.com/library/microcon-sx17-1003lqcef-1.

Biography

Mingyue Yao, M.Sc. in Electrical Engineering from Hong Kong University of Science and Technology. Currently she works as Engineer, Electrical and Control in FLAIR since 2020. Her research interests are artificial intelligence for production optimization and Industry 4.0.

Benny Drescher serves as Chief Technical Officer at FLAIR and Director Technology at INC, Hong Kong, SAR China. He holds a B.Sc./M.Sc. in Electrical Engineering and a Ph.D. (Dr.-Ing.) in Mechanical Engineering from Technical University of Munich, Germany.

Erik Stewart is Research Assistant in AI and Data Science at FLAIR. He received his B.S. in Industrial Engineering from Purdue University and has held various manufacturing-related engineering positions in the United States, Shanghai and Hong Kong.

Lakshmi Manoj is Research Assistant in AI and Data Science at FLAIR. She holds a Bachelor's Degree in Mechanical Engineering and works on implementing machine learning algorithms for industrial purposes.

Max Wittstamm received his M.Sc. in Production Engineering at RWTH Aachen University. Currently, he is Product Owner at FLAIR and works on the transfer of AI to production.

Leonhard Henke currently works as a Researcher and Project Manager at FIR in the field of information management. He holds a B.Sc./M.Sc. in Industrial Engineering from Otto von Guericke University of Magdeburg, Germany.

Shyam Ghodasara currently works as a Project Manager at the CIBA in the field of Business Software and its applications. He holds a M.Sc. in Management and Engineering in Production System from RWTH Aachen University, Germany.

Ming Ge, PhD. Expert in industrial AI and Robotics, works as Chief Operation Officer in FLAIR. He has rich experience in applied R&D on artificial intelligence, robotics, industrial big data, machine vision, model analysis and their applications for intelligent systems, intelligent manufacturing, automation, and so on.

3rd Conference on Production Systems and Logistics

Approach To A Decision Support Method For Feature Engineering Of A Classification Of Hydraulic Directional Control Valve Tests

Christian Neunzig[1,2], Simon Fahle[1], Jürgen Schulz[2], Matthias Möller[2], Bernd Kuhlenkötter[1]
[1]Chair of of Production Systems, Ruhr-University Bochum , Universitätsstraße 150 , 44801 Bochum, Germany
[2]Bosch Rexroth AG, Bexbacher Straße 72; 66424 Homburg, Germany

Abstract

Advancing digitalization and high computing power are drivers for the progressive use of machine learning (ML) methods on manufacturing data. Using ML for predictive quality control of product characteristics contributes to preventing defects and streamlining future manufacturing processes. Challenging decisions must be made before implementing ML applications. Production environments are dynamic systems whose boundary conditions change continuously. Accordingly, it requires extensive feature engineering of the volatile database to guarantee high generalizability of the prediction model. Thus, all following sections of the ML pipeline can be optimized based on a cleaned database. Various ML methods such gradient boosting methods have achieved promising results in industrial hydraulic use cases so far. For every prediction model task, there is the challenge of making the right choice of which method is most appropriate and which hyperparameters achieve the best predictions. The goal of this work is to develop a method for selecting the best feature engineering methods and hyperparameter combination of a predictive model for a dataset with temporal variability that treats both as equivalent parameters and optimizes them simultaneously. The optimization is done via a workflow including a random search. By applying this method, a structured procedure for achieving significant leaps in performance metrics in the prediction of hydraulic test steps of directional valves is achieved.

Keywords

Predictive Quality; Machine Learning; Quality Control; Feature Engineering; Decision Support Method

1. Introduction

Rising computer capacities and increasing data availability are expanding the horizon for knowledge generation in production [1]. Politically and socially, increased requirements for sustainability and resource consumption are moving into the focus of companies [2]. More and more companies are perceiving digitization as an opportunity [3,4]. Advancing digitalization and high computing power are drivers for the progressive use of ML methods on production data. [5] One solution strategy for improving existing production systems is knowledge extraction from production data using ML [6]. Predictive quality describes the ability to make data-driven predictions of product- and process-related quality in the manufacture and use of physical products [7,8]. The use of ML for predictive quality control of product characteristics contributes to increasing the efficiency and sustainability of future manufacturing processes. [9] However, a success factor for a ML project is sufficient pre-sampling of production data concerning dynamically deviating boundary conditions during data generation and optimization of data quality. [10] The further sections of the ML pipeline can be optimized based on a cleaned database [11]. Various ML methods such as gradient boosting methods have so far achieved promising results in industrial hydraulic use cases [12].

DOI: https://doi.org/10.15488/12177

For each prediction model task, there is the challenge of making the right choice of which algorithm with corresponding hyperparameters (HP) and which method of feature engineering (FE) is most appropriate to achieve the best predictions. FE is essential to the ML pipeline, as the overall performance of the model is highly dependent on the available features and enables the incorporation of domain knowledge into the ML pipeline [13]. By generating high-quality features, the effectiveness of an ML pipeline can be increased many times over an identical pipeline without dedicated FE and increases the interpretability of the trained model [14].

This work focuses on supervised learning, a subtype of ML in which the target values, so-called labels, of given data $\boldsymbol{\mathcal{D}} = \{(x_i, y_i) \in \mathbb{R}^d \times \mathbb{R}\}$ are known [15]. A ML model is fed with a training data set $\boldsymbol{\mathcal{D}}_{train}$ and evaluated against the validation data set $\boldsymbol{\mathcal{D}}_{valid}$, where x_i is respectively the input with dimension d and y_i is the target [16]. For the given dataset $\boldsymbol{\mathcal{D}}$, the goal is to find the algorithm $A_j \in \boldsymbol{\mathcal{A}}$ from hypothesis space $\mathcal{A}$ with HP combination $\lambda_j \in \Lambda$ from HP space Λ that achieves the highest accuracy under the premise of highest possible generalizability. The dataset $\boldsymbol{\mathcal{D}}$ is only an extract from the population, thus the model attempts to achieve the most accurate representation of the output for a given input by approximating a transfer function [16,17].

In this paper, a decision-support method is developed for selecting appropriate FE techniques while choosing the best HP combination of a predictive model for a dataset with temporal variability and validated on a real industrial value stream of directional control valves at Bosch Rexroth. By applying this method, structured guidance is given and the automation of manual efforts for achieving significant leaps in the performance metrics of an extreme gradient boosting (XGB) classifier is achieved. The method provides a qualitative evaluation of different FE techniques in hydraulic directional valve manufacturing and achieves significantly more accurate predictions than sequential pipeline optimization as shown in the experiment section.

2. Related Work

The ML pipeline and the simultaneous selection of the prediction algorithm and HP are explained. The importance of Fs application in various hydraulic applications will be addressed.

2.1 Combined Algorithm Selection and HP optimization problem

Within the ML pipeline, the task blocks of data preparation, the FE, model generation and evaluation are run through sequentially [18]. HÜTTER ET AL. describe the problem of automatic optimization of ML pipelines, where the choice is made between the correct ML algorithms and corresponding categorical HP, as the combined algorithm selection and HP optimization problem (CASH), see equation (1). The loss function is minimized against all folds k of the cross-validation (CV) for the training data set $\boldsymbol{\mathcal{D}}_{train}$ with $\boldsymbol{\mathcal{D}}_{train}^{(i)} = \boldsymbol{\mathcal{D}}/\boldsymbol{\mathcal{D}}_{valid}^{(i)}$ from all models considered to limit the effect of overfitting and increasing the generalizability [11].

$$\boldsymbol{A}_{\lambda^*}^{*} \in \underset{A_j \in \mathcal{A},\ \lambda_j \in \Lambda}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(\boldsymbol{A}_{\lambda}^{(j)}, \boldsymbol{\mathcal{D}}_{train}^{(i)}, \boldsymbol{\mathcal{D}}_{valid}^{(i)}) \quad (1)$$

ZÖLLER ET AL. extend the CASH problem to optimize the pipeline $P \in \mathcal{P}$ with the pipeline structure $g \in \mathcal{G}$ from a set of valid pipeline structures $\mathcal{G}$, where $|g|$ is the length of the pipeline, see (2) [14].

$$(g, \boldsymbol{A}, \boldsymbol{\lambda})^{*} \in \underset{A_j \in \mathcal{A}^{|g|},\ \lambda_j \in \Lambda}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(P_{g,\boldsymbol{A},\boldsymbol{\lambda},\boldsymbol{\mathcal{D}}_{train}^{(j)}}, \boldsymbol{\mathcal{D}}_{valid}^{(i)}) \quad (2)$$

2.2 Feature Engineering & CASH Optimization Methods for Hydraulic Use Cases

FE is about generating and selecting features from a given data set for the subsequent modelling step. FE can be split into three sub-tasks: feature extraction (FEX), feature construction (FC) and feature selection (FS) [19]. FS defines a subset of the feature set to speed up subsequent ML model training and improve its performance by removing redundant or misleading features. Simple domain-agnostic filtering approaches for FS are based on information theory and statistics. Methods such as univariate selection, variance threshold, feature importance, correlation matrices, or stability selection are integrated components of modern automated ML frameworks and are selected using standard CASH methods such as XGB algorithms [14]. Effective optimization approaches use random search or test the performance set exhaustively in a grid search also known as full factor design [16,20]. Advantages of random over grid search include easier parallelization and flexible resource allocation, since aborted iterations do not cause data holes and the combinations can be calculated independently [11]. Moreover, commonly used optimization methods contain reinforcement learning, evolution-based algorithm, and gradient descent, surrogate model-based optimization [18,21].

FS are using wrapper functions searches for the best feature subset by testing its performance on a given ML algorithm [19]. In heuristic approaches, individual features are added iteratively. Both forward and backward selection, as well as a combination of both, can be performed to select a subset of features. [18] In embedded methods, FS is directly integrated into the training process of an ML model. Many ML models, such as the Random Forest, provide some form of feature ranking that can be used. Similarly, embedded methods can be used in combination with FEX and FC [14]. Another optimization alternative is to use genetic programming in combination with prediction algorithms to identify a functioning feature subset [22]. TRAN ET AL. also used genetic programming to artificially construct novel features. In addition, information about how many times each feature was used during FC is reused to obtain feature importance [23]. KATZ ET AL. propose to compute meta-features for each novel feature, such as the diversity of values or the mutual information with the other features. Using a pre-trained classifier, the influence of a single feature can be predicted to select only promising features [24].

The prediction of internal leakage using discrete quality data has not yet been considered in science. The following are similar hydraulic industrial use cases whose approach has been considered in this work. LEI ET AL. recommend the combined use of principal component analysis (PCA) for dimensionality reduction of timeseries data with the classification and regression trees, random forests, and XGB, respectively, for the fault diagnosis model of the hydraulic valve [25]. KLUSCH ET AL. propose various statistical signal preprocessing steps such as correlation analysis and exploit this statistical characteristic in the form of features in a linear discriminant analysis and a k-nearest neighbor classifier for timeseries fault prediction [26]. HELWIG ET AL. identify times of equal boundary conditions in the prediction of hydraulic leakage flows and consider them during cross-validation and vary the weighting of the data within different time intervals in timeseries data [27,28].

3. Use Case: Internal Leakage Flow of Directional Control Valves

In this industrial use case at Bosch Rexroth, quality prediction of customer characteristics of directional control valves manufactured in Homburg is realized based on geometric gauge blocks from machining, mating data from assembly and hydraulic sensor data from end-of-line testing, shown in Figure 1. The target variable of this use case is the physically unavoidable internal leakage volume flow between spool and housing bore of a directional control valve [29]. The objective is subject to some uncertainty due to non-measurable variables and manufacturing tolerances. Previous work has attempted to understand the variables affecting leakage primarily through CFD simulations [30].

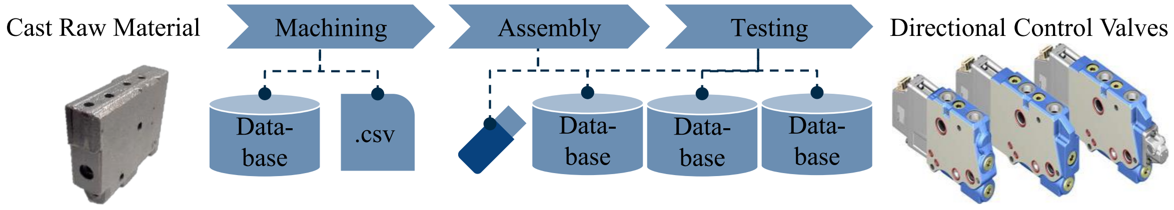


Figure 1: Value Chain of "Directional Control Valve" in Homburg.

An increased leakage volume flow leads to an unintentionally faster lowering of the load and involves dangers for people and property, see Figure 2a. The upper limit of leakage is secured in one of over 60 test steps as a safety-critical product characteristic within the scope of the final hydraulic inspection. In Figure 2b the 2-D view of the relevant valve area of the directional control valve is displayed. For the directional valve to function, the spool with different shoulders must be able to move within the housing bore, so that a ring gap between the bore and the spool, which is smaller than 20 micrometres, is inherent in the system, see Figure 2b [33]. The leakage measurement is carried out indirectly in the form of a pressure drop measurement at the pressure chamber of customer port A, which is reduced due to the leakage volume flow through the ring gap into the tank.

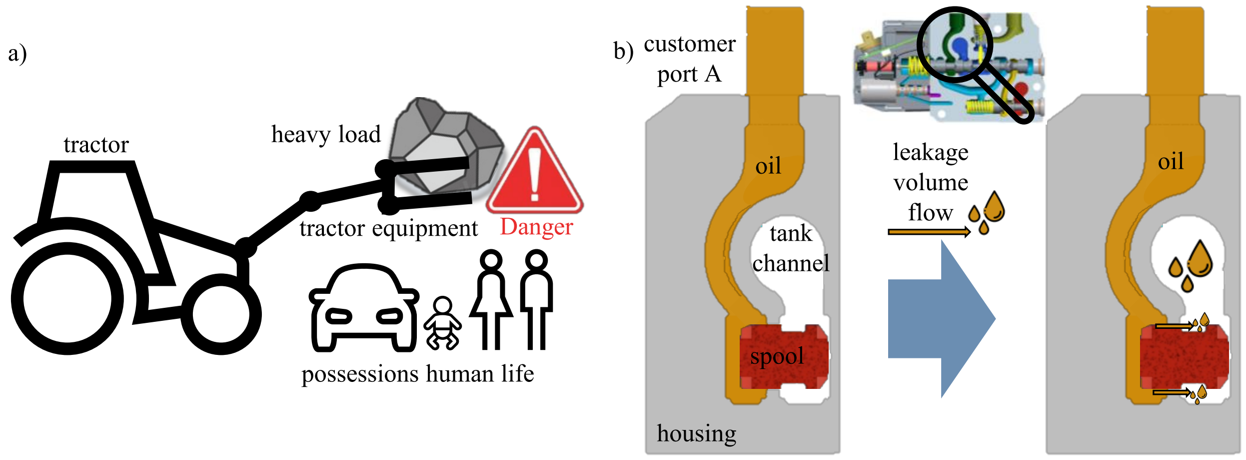


Figure 2: (a) safety-critical customer characteristic; (b) leakage volume flow as a pressure drop measurement.

4. Proposed Methodology & Experiments

The decision support method for selecting appropriate FE techniques is presented, applied to the hydraulic use case and finally the output is compared to a default as well as a tuned baseline model.

4.1 Decision Support Method for Selection of FE techniques & Pipeline Optimization

The proposed method for optimizing the ML pipeline and selecting the best fitting FE techniques builds on the CASH problem description of ZÖLLER ET AL [14]. The model matrix $\mathcal{M}$ forms the hypothesis space for the model candidates solving a prediction problem and consists of the three vectors each one for FE space $\mathcal{E}$, algorithm space $\mathcal{A}$ and HP space Λ, see Figure 3 and (4). The HP selection λ_w and partially the FE selection e_u is determined by the selection of the algorithm A_v since some algorithms will not run without appropriate FE. Thus, different types of algorithms require different numbers of FEs, so the matrix is the column dimension from the maximum of u, v, and w with u, v, w ∈ {1, 2, …, n}. The matrix is used to produce different combinations from the three vectors. Empty matrix cells contain a zero and are not considered as a combination.

$$\text{Model } \mathbf{M} \in \mathcal{M}_{3\,x\,n} = \begin{pmatrix} e_{u \in \{1, 2, \dots, n\}} \in \mathcal{E} \\ A_{v \in \{1, 2, \dots, n\}} \in \mathcal{A} \\ \lambda_{w \in \{1, 2, \dots, n\}} \in \Lambda \end{pmatrix} = \left(\begin{array}{cccc} e_1 & e_2 & \dots & e_n \\ A_1 & A_2 & \dots & A_n \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \end{array}\right) \begin{array}{l} \text{FE Space } \mathcal{E} \\ \text{Algorithm Space } \mathcal{A} \\ \text{HP Space } \Lambda \end{array}$$

Figure 3: Model Matrix of the Proposed Method for FE & Model Optimization.

The workflow of the method provides guidance on the structured approach to pipeline optimization and a qualitative trade-off of different FE techniques, see Figure 4.

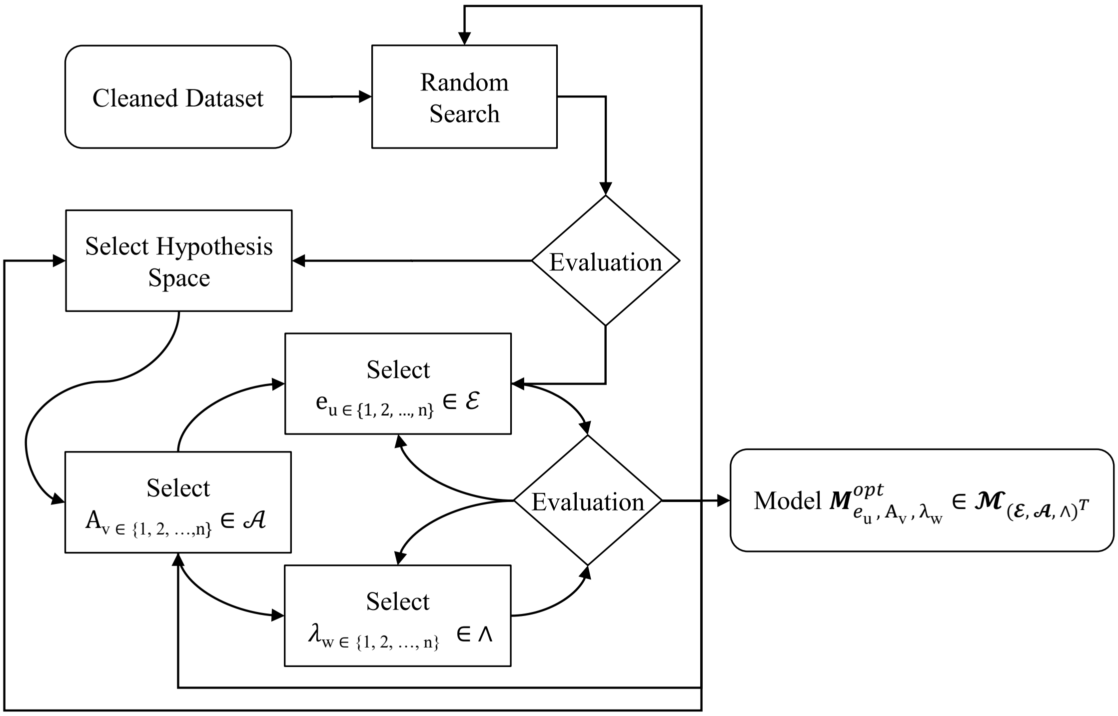


Figure 4: Flow Chart of the Proposed Method for FE & Model Optimization.

Various iteration loops illustrate the high combinability of the parameters. The random search is optimized with the given number of iterations. Based on a cleaned database, a random search is used to treat the FE techniques and the HP as equivalent optimization parameters. For the FE technique, this is partly achieved by providing a grid with Boolean data type. The inherent parameters of the FE techniques are treated as HP of the algorithms in the hypothesis space and are optimized according to the given grid. The procedure allows CV with k folds with different size as well for all combinations M$\in \mathcal{M}$. For optimization, the loss function from equation (4) is minimized against the validation data $\mathcal{D}^{(i)}_{valid}$.

$$\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w} \in \mathcal{M}_{(\mathcal{E}, \mathcal{A}, \Lambda)^T} = \boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w} \in \underset{\boldsymbol{M}^{u, v, w}_{e_u \in \mathcal{E}, A_v \in \mathcal{A}, \lambda_w \in \Lambda}}{\operatorname{argmin}} \frac{1}{k} \sum_{i=1}^{k} \mathcal{L}(\boldsymbol{M}^{(u,v,w)}_{\mathcal{E}_u, \mathcal{A}_v, \Lambda_w}, \mathcal{D}^{(i)}_{train}, \mathcal{D}^{(i)}_{valid}) \qquad (4)$$

4.2 FE Space: Derivation of used FE Techniques

Two core problems exist in the present production dataset: high dimensionality and high volatility over time. The production dataset consists of 11,652 data series with 1,052 features, which is a relatively unfavourable row-column ratio. The input features consist of geometric characteristics from machining, pairing information from assembly and sensor information from previous inspection steps. For this problem in a volatile high dimensional production data set, the XBG is a reasonable candidate of an algorithm [34]. The first task to be solved is the dimensionality reduction, see Figure 7. Preliminary studies have shown that

PCA is far superior to linear discriminant analysis just as forward selection is much better suited than backward selection for this data set. The final test data is composed as a matrix of the test steps (row-wise called PCA-A) and the sensors (column-wise called PCA-B). Second, the challenge of the changing boundary conditions of the production system manifest itself in value jumps in the data, as illustrated in Figure 5 for the scatter plots of two features. It is critical to recognize that the data jumps occur at different times for different feature columns. It is necessary to prevent the creation of new models each time the boundary condition changes, otherwise the amount of available data decreases drastically. Therefore, the change must be incorporated into the model by adjusting the same-sized fold CV. In future studies, identification with change point detection will be explored more intensively.

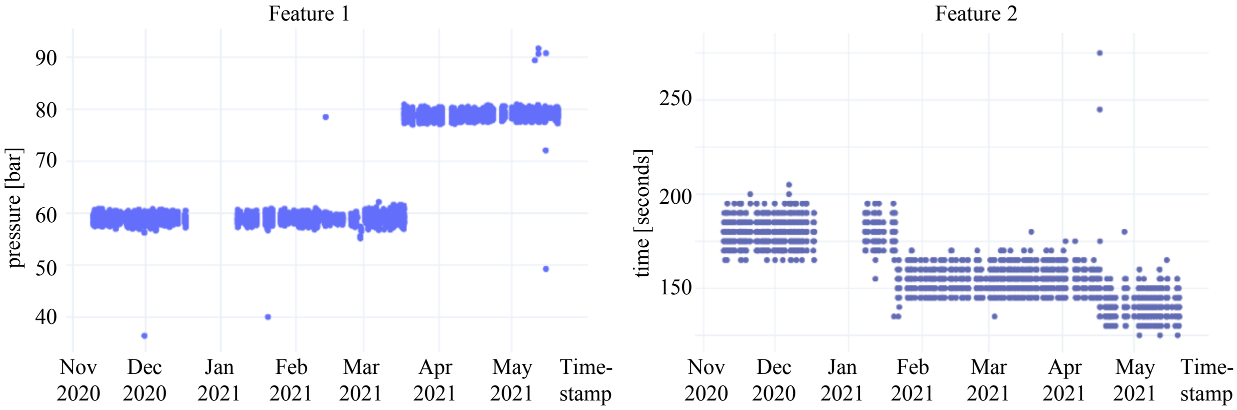


Figure 5: Exemplary scatter plots of two features with significant value jumps.

The 11 intervals with the same boundary conditions are indicated with the number of instances per interval, as shown in Figure 6. In this work, the interval-dependant timeseries 10-fold CV for each interval can be included in the presented workflow in Figure 4.

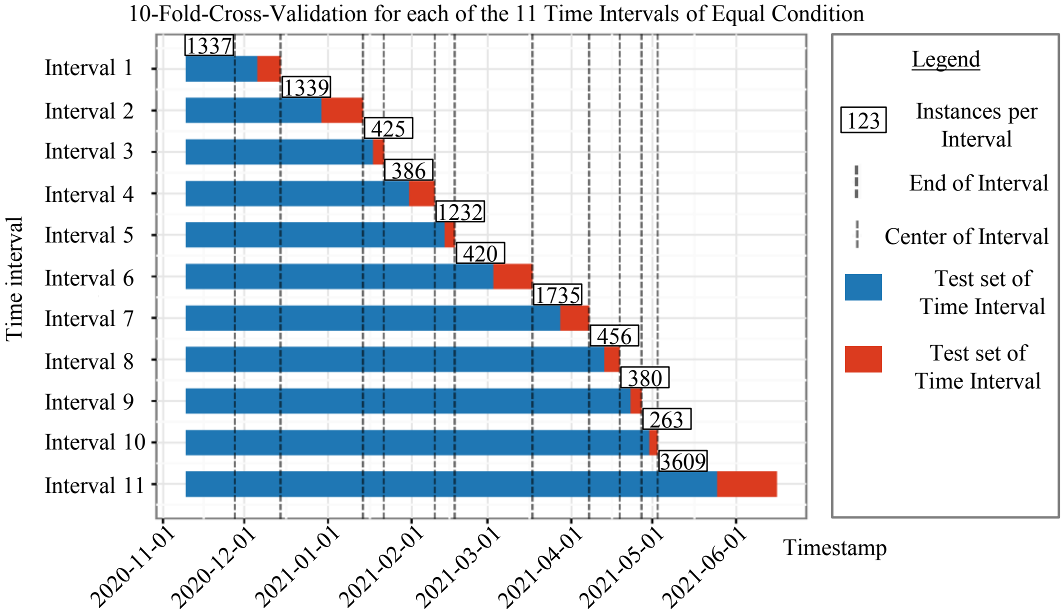


Figure 6: Timeseries-10-Fold-CV for folds for each time interval of equal systemic boundary conditions.

For the first interval, an 80/20 train-test split is performed. The following splits contain the previous splits and additionally the first half of the current split. The second half of the considered split is used as test data set. Within each interval, a 10-fold CV is performed and $\boldsymbol{\mathcal{D}}_{valid}^{(i)}$ is created. To account for the temporal variability of the data set, further weighting and consideration of different subsets are also applied using a window-based approach. More recent data points are weighted stronger within the loss function, applying exponential smoothing familiar from time series analysis and devaluing data points from earlier periods by a factor, using $\beta = (1-\beta)^k$ with $\beta \in \{0{,}1\}$ and k = 0 for the current interval and k = 1 for the following and so on, see Figure 7d. For the window-based approach, only the number n most recent data series backward to the prediction time are considered in the model to predict the state at the current prediction time, see Figure 7b. In addition, the data per interval are each transformed with centering and scaling, see Figure 7c.

All in all, in the proposed method a forward selection, a PCA depending on each test step (PCA-A) or sensor (PCA-B) are given as input for a dimensionality reduction and further compared to no application of dimensionality reduction at all. To evaluate the predictive power of data in the near and distant past of the prediction state, different amounts of data are utilized as input and a pre-optimized weighting function is applied. This weighting function is randomly activated and deactivated in the method, see Figure 7d. In addition, the data for all intervals is transformed to a comparable level via scaling and centering, see Figure 7d.

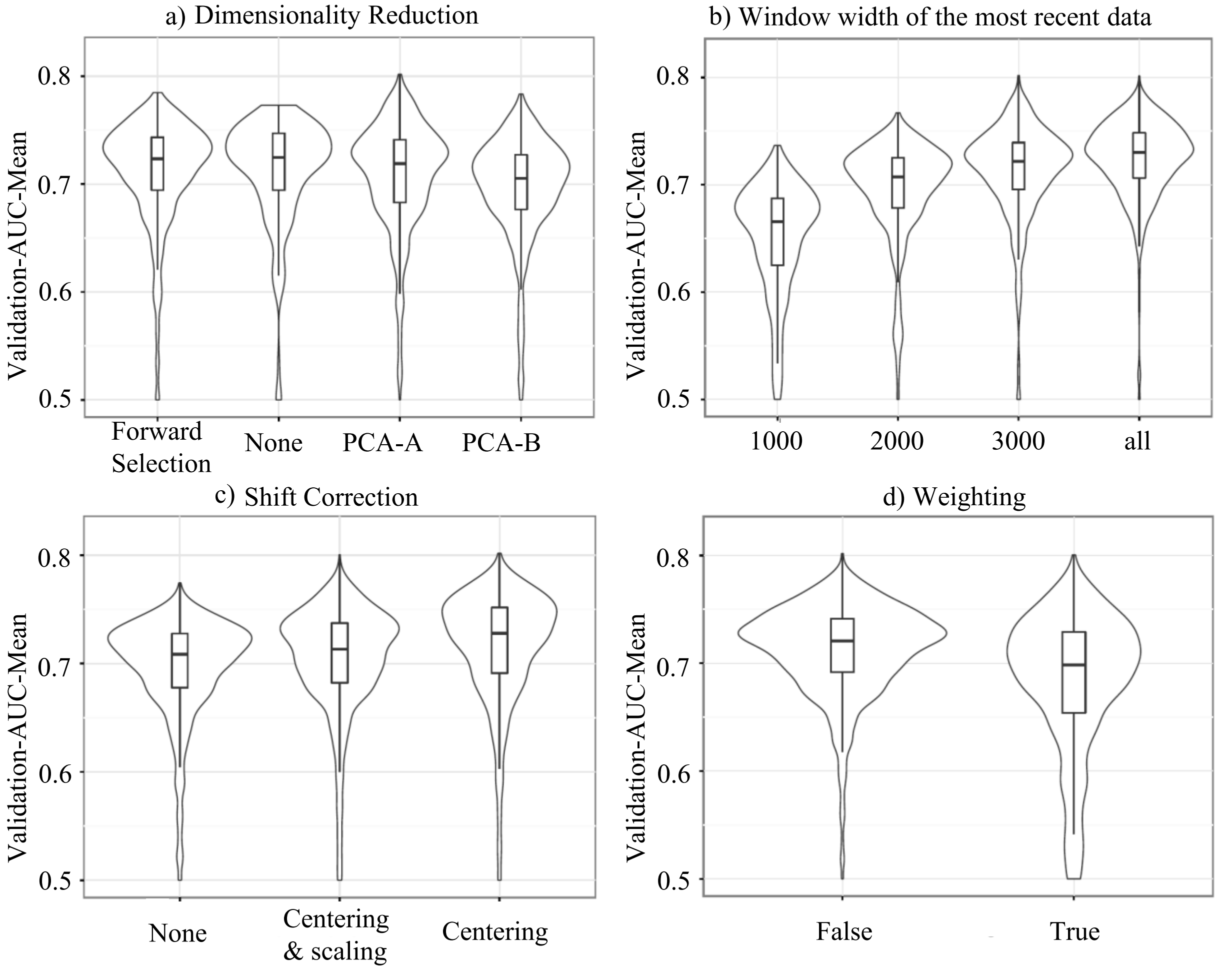


Figure 7: Qualitative Comparison of the FE techniques by Violin Plots.

4.3 Application and Evaluation of the Decision Support Method

A major strength of the presented method is the possibility of comparing several FE techniques while simultaneously applying and simultaneously optimizing the HP. Another advantage is the parallel calculation of the iterations within the method since the iterations are calculated independently. A suitable plot to compare the FE methods is the violin plot including frequencies and boxplot of AUC for different FE techniques, see Figure 7.

Thereby it turns out that the PCA-A with an AUC of 80 % produces the best candidate for the dimensionality reduction technique. Centering is the most effective technique for correcting the shift for this data set, as the best model shows the maximum AUC value of 80 % and averages 3 percentage points better than the variant with scaling. The weighting function should preferably be disabled in combination with listed FE techniques. Interesting to mention that the consideration of the last 3000 data series from the prediction time is not so much worse than the combination with consideration of the complete data set.

The proposed decision support method for FE techniques is applied to the hydraulic use case and compared with two baseline approaches to evaluate the suitability of the method, see Table 1. Many binary classification algorithms calculate the classification boundary and classify according to whether the value is above a certain threshold or not. The AUC is the trade-off between true-positive rate and false-positive rate that applies to all possible thresholds, not just the threshold chosen by the modelling technique. Different classification goals may result in one point on the curve being better suited for one task than another for different task. Hence, looking at the AUC is one way to evaluate the model regardless of the choice of a threshold. The F1-score brings precision and recall harmonized and is therefore very suitable as an evaluation of unbalanced data sets. The first base model (A) was applied to the data set with standard HP and without FE techniques. The second base model (B) consists of matched HP and preselected FE techniques, each of which produced the best results when applied individually. The application of centering, a weighting function with beta $\beta = 0.8$ and a PCA-B is performed for all data in model (B). The optimized model $\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w}$ is the best candidate of the proposed method after 5000 iterations with described FE input in 4.2, tuned HP of an XGB classifier and classifier itself. Sequential pipeline optimization and combination of preselected methods show a large effect on F1-score and AUC for unseen data. However, the application of the proposed method shows a significant jump in terms of F1-score by 10.4 percentage points and on the AUC of 8.3 percentage points for the unseen data compared to the sequential optimization.

Table 1: Comparison of different models for validation of the proposed method.

Model type	**FE**	**HP**	**AUC Train**	**AUC Test**	**F1-Score Train**	**F1-Score Test**
baseline model (A)	no methods	default	60,6 %	55,8 %	6.28 %	**3.33 %**
baseline model with FE (B)	pre-choice	tuned	68,7 %	**71,9 %**	11.3 %	**12.96 %**
optimized model $\boldsymbol{M}^{opt}_{e_u, A_v, \lambda_w}$	tuned	tuned	99,9 %	**80,2 %**	24.78 %	**23.36 %**

5. Conclusions and Future Work

The proposed decision support method produces the best combination of FE techniques and tuned HP for the XGB classifier on the time volatile production data of the industrial use case for predicting the leakage volume flow of directional control valves and outperforms the sequential optimization by about 10 percentage points in F1-score. From the qualitative review of various combinations of FE techniques, the combination of PCA-A, a data centring and a maximum window width with deactivated weighting emerged as the best FE combination.

Future work will validate a generalized suitability of the decision support method by applying it to additional algorithms and data sets. In addition, the optimization of the developed model and its implementation in production will be investigated. Feature construction approaches will be pursued to improve the usefulness of the model for series implementation by incorporating more predictive information into designed features.

References

[1] Posoldova, A., 2020. Machine Learning Pipelines: From Research to Production. IEEE Potentials 39 (6), pp. 38–42.

[2] Trauth, D., Bergs, T., Prinz, W., 2021. Monetarisierung von technischen Daten. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–734.

[3] Engels, B., Goecke, H., 2019. Big Data in Wirtschaft und Wissenschaft: Eine Bestandsaufnahme: IW-Analysen, No. 130. ECONSTOR, pp. 1–62.

[4] Lorenz, R., Benninghaus, C., Friedli, T., Netland, T.H., 2020. Digitization of manufacturing: the role of external search. IJOPM 40 (7/8), pp. 1129–1152.

[5] Krauß, J., Frye, M., Beck, G.T.D., Schmitt, R.H., 2019. Selection and Application of Machine Learning- Algorithms in Production Quality, in: Beyerer, J., Kühnert, C., Niggemann, O. (Eds.), Machine Learning for Cyber Physical Systems, vol. 9. Springer Berlin Heidelberg, pp. 46–57.

[6] Tao, F., Qi, Q., Liu, A., Kusiak, A., 2018. Data-driven smart manufacturing. Journal of Manufacturing Systems 48, pp. 157–169.

[7] Bergs, T., 2020. Internet of Production - Turning Data into Value, pp. 229–239.

[8] Wang, J., Ma, Y., Zhang, L., Gao, R.X., Wu, D., 2018. Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems 48, pp. 144–156.

[9] Krauß, J., Pacheco, B.M., Zang, H.M., Schmitt, R.H., 2020. Automated machine learning for predictive quality in production. Procedia CIRP 93, pp. 443–448.

[10] Klinkenberg, R., 2005. Meta-Learning, Model Selection, and Example Selection in Machine Learning Domains with Concept Drift. Lernen, Wissensentdeckung und Adaptivität (LWA), pp. 1–8.

[11] Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated Machine Learning. Springer International Publishing, Cham, pp. 84–95.

[12] Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., Wrobel, S., 2019. A review of machine learning for the optimization of production processes. International Journal Advanced Manufacturing Technology 104 (5-8), pp. 1889–1902.

[13] Chu, X., Ilyas, I.F., Krishnan, S., Wang, J., 2016. Data Cleaning, in: Özcan, Koutrika et al. (Hg.)–Proceedings of the 2016 International, pp. 2201–2206.

[14] Zöller, M.-A., Huber, M.F., 2021. Benchmark and Survey of Automated Machine Learning Frameworks. jair, pp. 409–472.

[15] Hastie, Trevor, Jerome Friedman und Robert Tibshirani, 2001. The Elements of Statistical Learning.: Data Mining, Inference, and Prediction, pp. 29–39.

[16] Feurer, M., Hutter, F., 2019. Automated Machine Learning: Hyperparameter Optimization. Springer International Publishing, Cham, pp. 7–33.

[17] Raschka, S., 2018. Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning, pp. 1–49.

[18] He, X., Zhao, K., Chu, X., 2021. AutoML: A survey of the state-of-the-art. Knowledge-Based Systems 212, pp. 1–37.

[19] Motoda, H., Liu, H., 2002. Feature Selection Extraction and Construction. Communication of IICM (Institute of Information and Computing Machinery, Taiwan) Vol 5., pp. 67–72.

[20] Montgomery, D.C., 2013. Design and analysis of experiments, John Wiley & Sons. pp. 1–757.

[21] Nicolo Fusi, Rishit Sheth, Melih Elibol, 2018. Probabilistic Matrix Factorization for Automated Machine Learning. Neural Information Processing Systems, Montréal, Canada.

[22] Olson, R.S., Moore, J.H., 2019. TPOT: A Tree-Based Pipeline Optimization Tool for Automating Machine Learning, in: Hutter, F., Kotthoff, L., Vanschoren, J. (Eds.), Automated Machine Learning. Springer International Publishing, Cham, pp. 151–160.

[23] Tran, B., Xue, B., Zhang, M., 2016. Genetic programming for feature construction and selection in classification on high-dimensional data. Memetic Comp. 8 (1), pp. 3–15.

[24] Katz, G., Shin, E.C.R., Song, D., 2016. ExploreKit: Automatic Feature Generation and Selection, in: 2016 IEEE 16th International Conference on Data Mining (ICDM). 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain. 2016. pp. 979–984.

[25] Lei, Y., Jiang, W., Jiang, A., Zhu, Y., Niu, H., Zhang, S., 2019. Fault Diagnosis Method for Hydraulic Directional Valves Integrating PCA and XGBoost. Processes 7 (9), pp. 589 –607.

[26] Klusch, M., Meshram, A., Schuetze, A., Helwig, N., 2015. iCM-Hydraulic, in: Proceedings of the 11th International Conference on Semantic Systems. SEMANTiCS '15: 11th International Conference on Semantic Systems, Vienna Austria. 2015. ACM, New York, NY, USA, pp. 81–88.

[27] Helwig, N., 2018. Zustandsbewertung industrieller Maschinen mittels multivariater Sensordatenanalyse. Dissertation, pp. 1–210.

[28] Helwig, N., Pignanelli, E., Schutze, A. Condition monitoring of a complex hydraulic system using multivariate statistics, in: 2015 IEEE International Instrumentation, pp. 210–215.

[29] Matthies, H.J., Renius, K.T., 2014. Einführung in die Ölhydraulik. Springer Fachmedien Wiesbaden, pp. 38–48.

[30] Posa, A., Oresta, P., Lippolis, A., 2013. Analysis of a directional hydraulic valve by a Direct Numerical Simulation using an immersed-boundary method. Energy Conversion and Management 65, pp. 497–506.

[31] Morris, N., Rahmani, R., Rahnejat, H., King, P.D., Fitzsimons, B., 2013. Tribology of piston compression ring conjunction under transient thermal mixed regime of lubrication. Tribology International 59, pp. 248–258.

[32] Neunzig, C., Fahle, S., Kuhlenkötter, B., Möller, M., 2021. Feature Engineering For A Cross-process Quality Prediction Of An End-of-line Hydraulic Leakage Test Using An Experiment Sample, 2nd Conference on Production Systems and Logistics, pp. 156–166.

[33] Amirante, R., Del Vescovo, G., Lippolis, A., 2006. Flow forces analysis of an open center hydraulic directional control valve sliding spool. Energy Conversion and Management 47 (1), pp. 114–131.

[34] Mangal, A., Kumar, N., 2016. Using big data to enhance the bosch production line performance: A Kaggle challenge., in: 2016 IEEE International Conference, pp. 2029–2035.

Biography

Christian Neunzig has been a member in the Bosch PhD program since 2019. Previously, he completed a master's degree in production engineering and a bachelor's degree in mechanical engineering at RWTH Aachen University. His research focus is applied machine learning on directional control valve manufacturing data.

Simon Fahle has been a member of the Chair of Production Systems (LPS) at Ruhr University Bochum since 2019. He completed a bachelor's and master's degree in mechanical engineering at Ruhr-University Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Jürgen Schulz is manager of the process planning group and assembly and testing technology at the Bosch Rexroth plant in Homburg. Mr. Schulz received diploma (Dipl.-Ing.) from the University of Applied Sciences in Saarbrucken (HTW).

Matthias Möller is head of the Technical Functions department at the Bosch Rexroth plant in Homburg. Mr. Möller received his doctorate (Dr.-Ing.) and diploma (Dipl.-Ing.) from the Technical University of Dortmund.

Bernd Kuhlenkötter was responsible for product management and technology at ABB Robotics Germany until 2009. In 2009, Bernd Kuhlenkötter took over the professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter has held the professorship for "Production Systems" at the Ruhr University Bochum since 2015.

3rd Conference on Production Systems and Logistics

Towards A Data-driven Performance Management In Digital Shop Floor Management

Lukas Longard[1], Sebastian Bardy[1], Joachim Metternich[1]
[1] *TU Darmstadt, Institute of Production Management, Technology and Machine Tools (PTW), Darmstadt, Germany*

Abstract

Key performance indicators (KPIs) are crucial for measuring and managing the performance of industrial processes. They are used to detect deviations in processes, enabling opportunities to improve manufacturing processes within the three dimensions time, quality, and cost.

In this context, the timeliness of information plays a decisive role in the success of measures since delayed information availability can leave decision makers with no time to react. With the introduction of digitization and industry 4.0, increasing amounts of data become available. They can be used to accelerate problem detection and shortening reaction times to define appropriate actions.

This paper presents a data-driven performance management approach integrated in digital shop floor management (dSFM). If a deviation is detected in one process, KPIs of subsequent processes (horizontal level) as well as subordinate levels (vertical level) are checked for correlations and, if present, the associated team is notified by an automatic warning through the dSFM system. Based on the identified correlations, the team discusses the deviations and defines suitable countermeasures. The aim of this approach is to identify deviations more quickly and to quantify their impacts, thus giving shop floor managers the ability to react in time.

Keywords

Shop floor management; Performance management; Key performance indicators; Data Mining; Machine Learning

1. Introduction

Recent advances in digitization offer a high potential for companies operating in the manufacturing domain to reduce reaction time on business-relevant events like unplanned downtimes and quality issues [1]. Providing the right information to the right people at the right time in an efficient manner to empower them to make the right decisions and take the right course of actions is a significant difficulty for many producing organisations [2,3]. If this can be done in a timely fashion, the negative effects of deviations can be reduced and impacts on internal or even worse on external customers can be prevented [1].

The methods of shop floor management (SFM) are widely used in industry to control and improve production processes on a daily basis [4]. One of the most important elements of SFM is performance management. To manage process performance, goals are set by the management and translated into trackable key performance indicators (KPIs) to identify deviations in processes [5]. These are then analysed in shop floor meetings and a problem-solving process is initiated if necessary. Improvements developed in the problem-solving process are stabilized and standardized to reach a continuous improvement of the production processes [6,7]. However, there are several shortcomings of performance management and its application in industry.

DOI: https://doi.org/10.15488/12185

Hellebrandt et al. state that performance management is mainly used in middle and top management and KPIs on the shop floor are not connected to these higher levels [8,9]. Furthermore, KPIs on the lowest level are not connected to the individual worker, making it difficult to achieve a sense of responsibility by the employee towards the KPIs [10]. Moreover, due to the large number of KPIs often used, the complex interrelationships can no longer be intuitively understood and anticipated, resulting in a great demand for system-based decision support [11].

Therefore, this paper will present a new data-driven performance management approach in digital SFM (dSFM). The remainder of the paper is structured as follows: Chapter 2 provides the state of the art on SFM, performance management as well as recent advances. Chapter 3 introduces the model of latency to business-relevant events and derives the goals and opportunities of a data-driven performance management approach. Following up, the data-driven approach is described in chapter 4. Finally, the paper closes with a conclusion and outlook for the next steps in the development.

2. State of the art

2.1 Shop floor management and performance management

Hertle et al developed a model to describe the daily routine for a successful SFM (see Figure 1). Based on standardised processes, production goals are set by management. In step one, deviations from the set goals are identified with the help of target-actual comparisons of KPIs, andon or gemba walks. In step two, the deviations are discussed in daily shop floor meetings. The impact of the deviation is evaluated, and short-term countermeasures are initiated. A decision is also made as to whether a systematic problem-solving process (SPSP) should be started. The SPSP is not part of the daily routine and runs separately. A PDCA cycle is used to track the progress of implementation. Step three of the SFM loop comprises the first two phases (Plan & Do) of the PDCA cycle. In the final step, the measures introduced are checked and tested for suitability so that they can be transferred to the standard in the event of a positive vote [5].

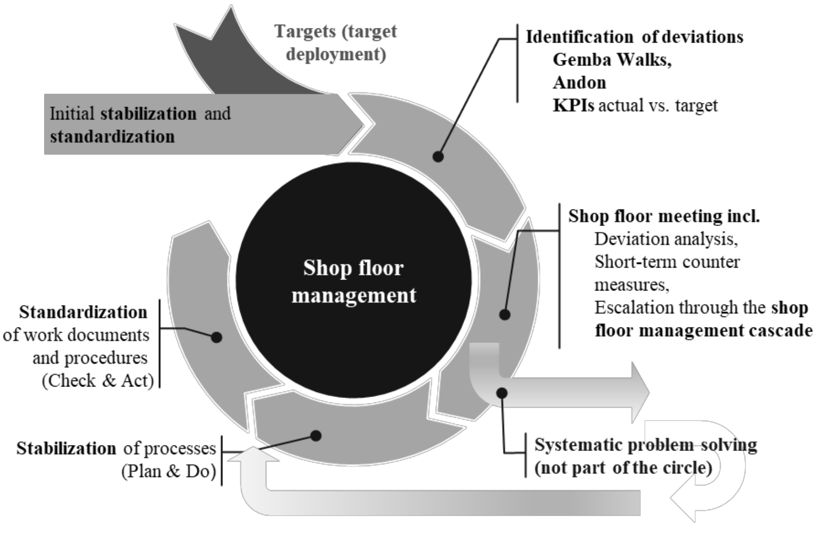


Figure 1: Shop floor management model [5]

To implement KPIs there are two main prerequisites. Firstly, management must define targets for the production processes and, secondly, ensure multidimensional measurement of production performance (performance measurement) in order to be able to visualize the degree of target achievement [3,12]. If every target is linked to an improvement activity which supports the achievement of the long-term vision of the company, the approach is called Hoshin Kanri [13]. In this context a performance pyramid is often used for visualization (see Figure 2). Based on the corporate vision, strategic goals are derived for the three performance levels of strategical management, tactical management, and operational level in the sense of a top-down approach. The achievement of the goals in the respective levels is determined by KPIs. The indicators are aggregated in a bottom-up approach so that causal relationships exist between the indicators of the different levels. [12,13]

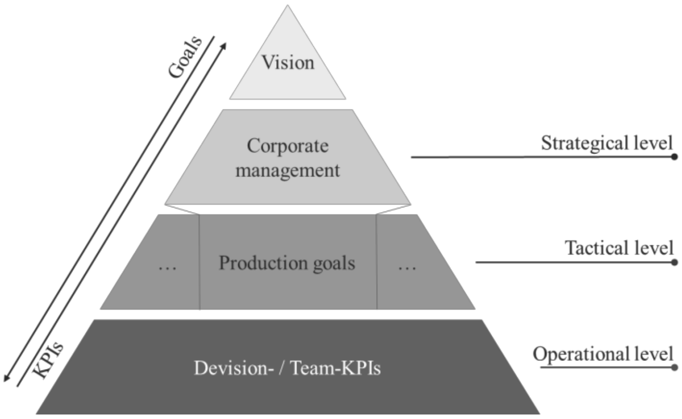


Figure 2: Performance pyramid for production, adapted from [14]

2.2 Recent advances in shop floor management, performance management and problem solving

With the introduction of digitization and industry 4.0, increasing amounts of data become available for processing and use in smart manufacturing systems [15]. Meissner et. al developed a target state for dSFM. They suggest using KPI data to forecast KPIs and predict trend impacts of upstream or downstream processes.. Then the information is visualized and managers of the process as well as downstream processes are warned. [16] By integrating machine and manufacturing data into the performance pyramid, KPI calculation can be automated [17] and generated real-time data enable further insights [18].

The new possibilities in dSFM are not only viable for performance management but also translate to problem management. In classic problem management, the deviations in KPIs are presented to the employees/managers where they must make the decision how to handle the deviation. This can be categorized into three different levels of reaction: If the deviation is not impactful or even a false alarm it can be ignored. If the deviation has an impact on production performance and the root cause is clear immediate action should be taken to prevent further losses. Finally, if the deviation has an significant impact on target fulfilment and cause is unknown a systematic problem-solving process (SPSP) is used to find the right countermeasure. [19]

However, classic detection mechanisms like KPIs are often only able to detect the symptoms of underlying problems. Remedying those symptoms is not sufficient to resolve the underlying problem and to find a sustainable solution [20]. Without a systematic approach to problem-solving, employees are tempted to hastily identify causes and introduce immediate measures. These are usually based on experience and feelings, but not on a sound analysis of the root cause of the problem at hand. German studies have shown that up to 60% of emerging problems are recurring [21], which indicates that it is rare that lessons are learned from past mistakes and the root cause of problems is sustainably eliminated [22]. Meissner et al. put in perspective that digitalization can enrich the information available for root-cause analysis. Furthermore, through algorithms root-causes as well as solutions for the problems can be proposed by the system to the employee. [23] To comprehend these complex relationships, data mining (DM) can be used as an analysis support [24]. In their literature review, Longard et. al show the potentials of using DM in SPSP. As problem solving requires a lot of experience and creativity, humans are superior to machines and computers in this field. Data can especially support hypothesis formulation and problem delimitation as well as analysis. In particular, correlation analyses between miscellaneous process parameters can provide valuable insights to support the interpretation of the results and prepare the creative work in finding solutions. [25]

3. Goals of a data-driven performance management

Hackarthon developed a model for business intelligence that considers the different time elements between the occurrence of a business-relevant event and the initiation of remedial action (reaction time) that can be transferred to the domain of SFM. According to the model, the reaction time can be decomposed into data, analysis, and decision latency [26]. The longer the process takes from the occurrence of a business-relevant event, through detection and analysis, to the initiation and implementation of countermeasures, the more business value is lost (see Figure 3 - left).

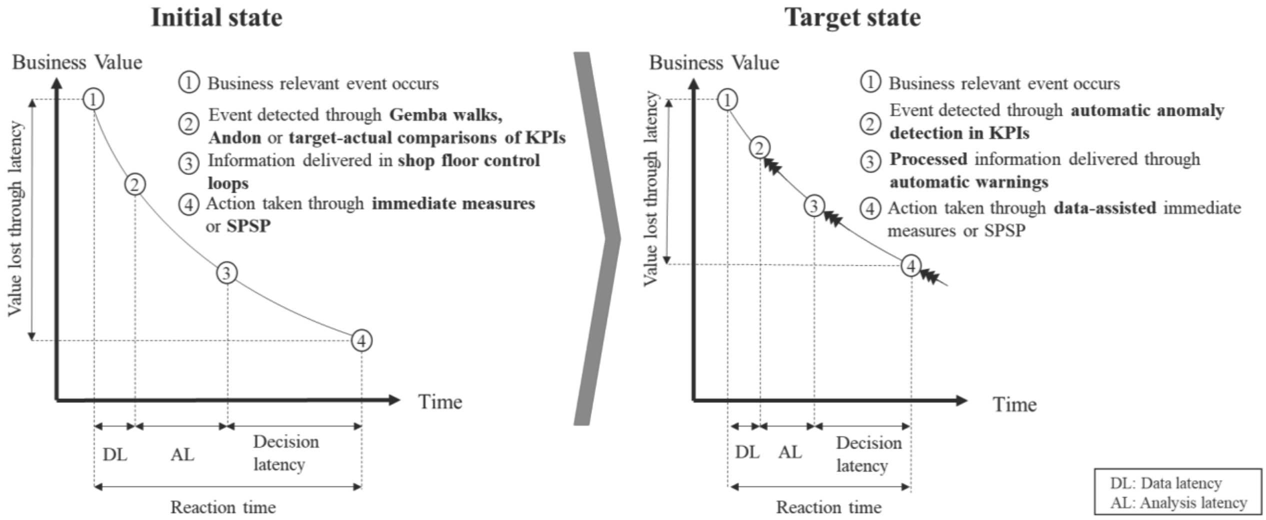


Figure 3: Deviation management in SFM – initial state versus target state, adapted from [1,26]

The data or detection latency describes the delay between the occurrence of a fault and its capturing. In the classical sense, this is recognized in SFM by means of Gemba walks, the target-actual comparison of KPI or Andon signals triggered by employees as well as machines [5]. These simple methods are able to identify many process or product deviations in order to restore the desired condition. Nevertheless, valuable time is lost since these measures only have a delayed effect on the actual cause and are therefore considered reactive measures. Even though the automated calculation of KPIs is an important step towards reducing detection latency, the gap between the occurrence and detection of a business-relevant event can only be closed by connecting sensors that measure as close as possible to the actual root cause. The described relationships are shown in Figure 3. The use of automatically calculated KPIs and sensor data can lead to a reduction in the detection latency (shift upwards along the curve).

Reducing detection latency to a minimum only has a positive effect if the decision-makers receive the relevant information in time [27]. Zur Mühlen et al. define the analysis latency as "[…] the delay between the storage of event information in a repository and the subsequent transformation of this event information into an analysable format, such as a notification, report, or indicator value." [1]. This is where traditional SFM systems with their fixed communication cycles [7] and rudimental information (e.g. visualization of KPIs) [8] reach their limits and therefore have to be adjusted. It is particularly important to quantify impacts of deviations on subsequent processes as well as subordinate levels. Moreover, to exploit the full potential of the data, decision-makers must receive information on relevant events as quickly as possible and in a form that is easy to understand. Especially, when dealing with sensor data, without contextual information, it is almost impossible to evaluate a situation and draw the right conclusions [28]. In addition, the right amount of information has to be determined to not cause an information overload [3].

After detecting (e.g. through anomaly detection on sensor data) and transforming the information into an analysable format, adequate remedial actions have to be initiated. Decisions must be made quickly, and the decision latency must be kept as low as possible to minimize the impact on business value (see Figure 3). In contrast, the root cause of a problem and not just its symptoms should be addressed through SPSP to benefit in the long run. The use of immediate measures should therefore only be used for damage limitation and should not replace a SPSP. Both the selection of an immediate measure and the root cause analysis with the underlying cause-effect relationships require in-depth knowledge.

In summary, to reduce the reaction time to business-critical events and minimize resulting value losses, a data-driven performance management approach in dSFM must address the following shortcomings of current approaches:

- **Goal 1**: To be able to recognize deviations earlier, information must be available as quickly as possible. Data (especially from sensors) should be used to shorten the gap between occurrence and detection of business-relevant events.
- **Goal 2**: Decision-makers should receive information on relevant events as quickly as possible in the right amount and quality.
- **Goal 3**: To enable prioritization, the impact of deviations on subsequent processes or higher levels should be quantified.
- **Goal 4**: Data should assist problem solvers in finding the root causes faster, thus shortening the decision latency.

4. A data-driven performance management approach

The developed approach aims to quantify the potential impact of business-critical deviations at the horizontal and vertical level, alert the associated operations managers, and give them time and information to define appropriate countermeasures. Here, the horizontal level refers to the value stream and attempts to quantify the effects of deviations on subsequent process steps. This is to enable the subsequent processes to react to the impending effects and to take appropriate measures. If a deviation is detected in one process, KPIs of subsequent processes are checked for (time-lagged) correlations and, if present, the associated team is notified by an automatic warning through the dSFM system. Based on the warnings, the team discusses the deviations and defines suitable countermeasures (see Figure 4 - left). In contrast, the vertical level describes the effects of a deviation in a process along the company hierarchy. If again a deviation is detected in one process, KPIs of higher-levels are checked for correlations and, if present, the ones responsible are notified (see Figure 4 - right). The objective is to quantify the impact of sub-areas at aggregate levels to inform higher-level managers when production goals are in jeopardy. This is intended to simplify the escalation process and give quantitative reference.

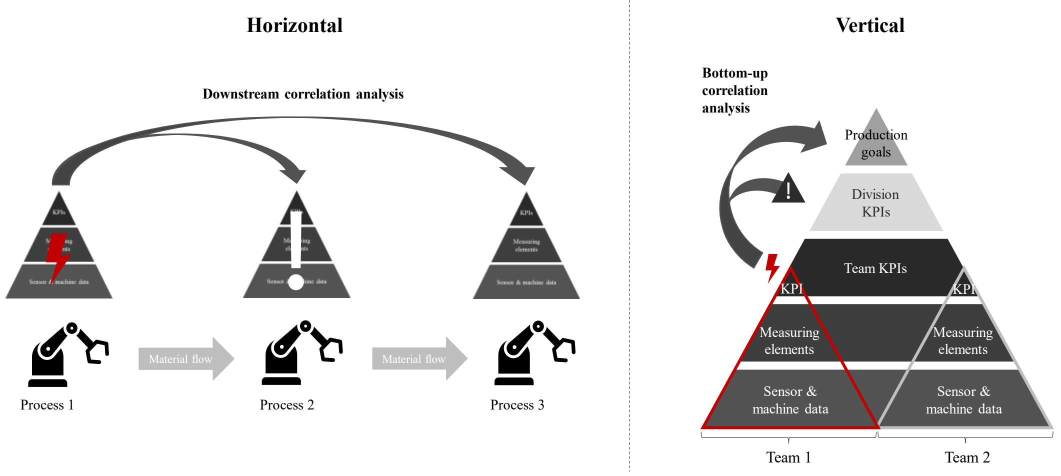


Figure 4: A data-driven performance management approach

In comparison to traditional performance management approaches, which focus primarily on the vertical consistency of business goals using KPIs, this new approach is intended to form a KPI network that also considers horizontal dependencies to promote value-stream-wide collaboration. The fact that these KPIs represent all 3 dimensions - time, quality and costs - means that a targeted focus for improvement can be established. To implement this data-driven performance management approach, the three latencies (data, analysis & decision) are addressed in a targeted manner through the three phases of deviation detection, impact quantification as well as warning & impact assessment (see Figure 5). These will be discussed in the following.

4.1 Deviation detection

The starting point of the SFM control loop is the detection of deviations [16,23]. As described in Chapter 3, current approaches are not able to fulfil the requirement of using data to bridge the gap between the occurrence and detection of business-relevant events (Goal 1 + 4). For this reason, a three-level approach was defined that starts at a high level with anomalies in KPIs and gradually gets finer by incorporating their measurement elements up to machine/sensor level data (see Figure 5 – Deviation detection). From a technical point of view, the detection of deviations requires different methods and algorithms for each level. On the KPI-level a simple target-actual comparison realized by a corridor with upper and lower limit is sufficient to capture most of the relevant deviations. Since KPIs are often calculated from a large number of so-called measurement elements (e.g. good quantity, part quantity, actual unit processing times), a deviation detection only at KPI level would lead to a certain lack of clarity and make root cause analyses more difficult. Therefore, the next step is to look at this level. The time series of the measurement elements have similar properties to the KPIs with the difference that higher measurement frequencies are often available. This is due to the fact that KPIs are often formed only once per shift or day, but the underlying measurement elements are recorded more frequently and are thus available for analysis. In contrast to detecting KPI deviations, applying target-actual comparisons on the measuring elements is not applicable, since there are usually no specified targets for those. One way to solve this problem is to define dynamic target values (e.g. dependence on time and product). In addition, statistical process control and trend analysis, could provide valuable results.

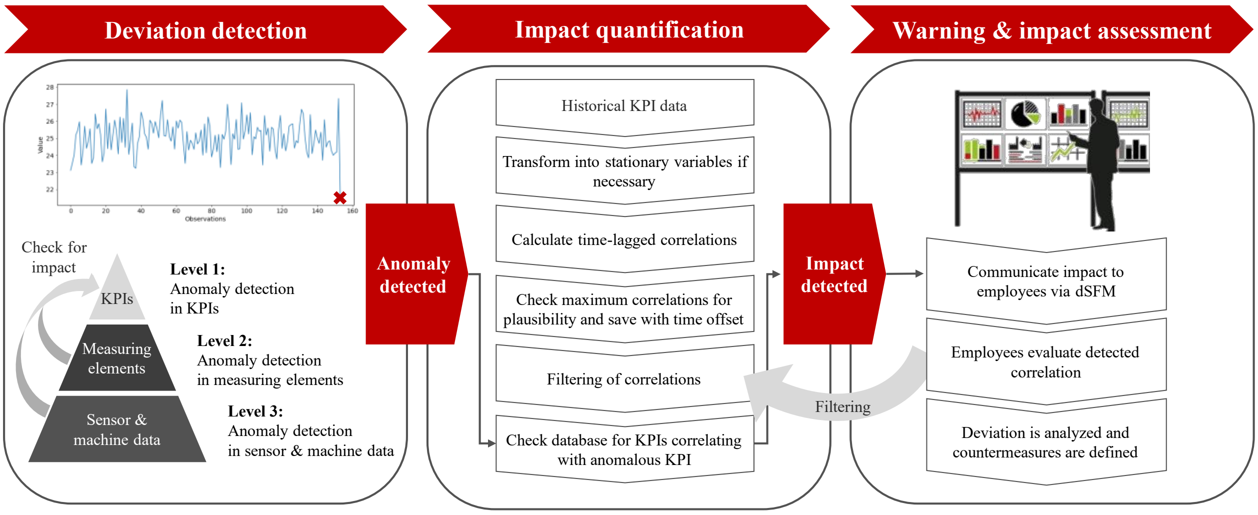


Figure 5: Model pipeline of data-driven performance management approach in dSFM

With the integration of the sensor and machine data, the goal is to measure as close as possible to the root cause of a problem or deviation. The integration of new measuring points provides the possibility to get even closer to the process, which is not yet done today. To find anomalies in the high-frequency data of sensors or machines, it is recommended to use more sophisticated methods and algorithms. The rapid developments in this direction in recent years have produced many such methods and algorithms. These include machine learning (ML) and artificial intelligence (AI) approaches like support vector machines [29] and neural networks [30] but also statistical approaches like ARMA or ARIMA [31] to name a few.

4.2 Impact quantification

The knowledge of the quantitative impact of deviations can only be generated from long-term data. This data must be pre-processed and transformed into stationary variables, to reduce the probability of encountering spurious correlations. In the next step, the relationships between the different KPIs are quantified via (time-lagged) correlation analyses. The focus on time-lagged KPIs is due to the fact that it is precisely those impacts that are interesting from a management perspective, which have a time-lagged reaction and can thus be counteracted by an action (Goal 2 + 3). In addition, the direction of the correlation must be determined to be able to make a statement about positive or negative impact of the leading KPI on the lagging KPI. Thereby, the model must also reflect domain knowledge, since different KPIs have different optimization goals (e.g., maximizing OEE as opposed to minimizing scrap rates). In order to be able to capture the multitude of different correlations (e.g. linear, quadratic) between KPIs, more advanced methods must be used in addition to the standard correlation methods Pearson, Spearman, Kendall (only able to capture linear correlations). In the recent past, the Maximum Information Coefficient has stood out and will be taken into account in future studies [32]. After the maximum lagged correlation has been determined for each pair of KPIs, the determined offset must be checked for plausibility. For example, from a practical point of view, it may not make sense if the detected offset is larger than the lead time between the processes belonging to the KPIs. Afterwards, all detected correlations and their corresponding offsets are saved in a database.

The final step of the impact quantification phase is to match detected anomalies with the detected correlations. If an anomaly is detected in a KPI, the database is searched and correlations belonging to the KPI are returned. If an anomaly is detected at the measuring element or sensor level, it is first checked (e.g. by correlation or regression analyses) whether this has an impact on the KPIs of the associated process (see Figure 5 – Deviation detection). If this can be confirmed, the procedure is the same as described above.

4.3 Warning and impact assessment

The next step is to notify those managers whose KPIs correlate with the anomalous KPI. To keep the latency as low as possible, it is advisable to send the warnings via mobile devices, emails or push messages in the dSFM. The criticality of the deviation should be used when choosing the communication medium. This can be determined by an interaction of the correlation coefficient, the temporal offset, possible effects on higher levels, and employee-defined intervention limits and assessments of past cases. To make the information processable for the employees, it must be prepared in a suitable form (Goal 2). This can be achieved both by the form of visualization and by context provided for the information [33]. This includes information on when the impact is likely to occur, which of the team´s own KPIs are affected, and which KPI (which team) is the cause of the deviation. In addition, context is also given to similar warnings that have occurred in the past. After that, the employees evaluate the warning based on the available information. In doing so, they are given the opportunity to evaluate the correlations recognized by the algorithm, for example, to hide spurious correlations for future warnings. In this way, the underlying model is continuously improved by the employee (active learning). Finally, a decision is made as to whether an action or SPSP should be initiated or whether the information should merely be noted and communicated to employees in the dSFM.

5. Conclusion and outlook

In this publication, a data-driven performance management approach for dSFM is presented, focusing on the three steps deviation detection, impact quantification and warning & impact assessment. The goal of the approach is to significantly reduce the time between the occurrence of a business-relevant event and the initiation of remedial action to prevent the loss of business value. To achieve this, the concept of three latencies, data latency, analysis latency and decision latency was introduced and countermeasures for reduction were developed. At the core of the approach is anomaly detection at the KPI, measurement element, and sensor/machine level using ML and AI algorithms and quantifying the impact of these anomalies on downstream processes as well as higher hierarchical levels through correlation analyses.

After the data-driven performance management approach in dSFM has been developed in this paper, a practical evaluation of the individual phases will be carried out in the future. To achieve this, a dSFM system available on the market will be further developed around the data-driven performance management approach and put into real use at a company from the process industry. In particular, it will be investigated which different correlation methods are suitable for quantifying the effects and how these correlations can be prefiltered automatically (e.g., from spurious correlations). To not only uncover that a relationship exists (correlation), but also to quantify the magnitude of that relationship, regression models for KPIs will be built in the future. Furthermore, from a research point of view, it will be interesting to see whether the described approach can increase production performance and what factor the integration of sensor and machine data plays.

Acknowledgements

This project (HA project no. 1012/21-14) is financed with funds of LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientific and Economic Excellence).

References

[1] zur Mühlen, M., Shapiro, R., 2010. Business Process Analytics, in: vom Brocke, J., Rosemann, M. (Eds.), Handbook on Business Process Management 2. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 137–157.

[2] Harris, C., 2012. Lean Connections: Making Information Flow Efficiently and Effectively. Taylor and Francis, Hoboken, 166 pp.

[3] Eaidgah, Y., Maki, A.A., Kurczewski, K., Abdekhodaee, A., 2016. Visual management, performance management and continuous improvement. International Journal of Lean Six Sigma 7 (2), 187–210.

[4] Pötters, P., Schindler, P., Leyendecker, B., 2018. Status quo Shopfloor Management. Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (7-8), 522–525.

[5] Hertle, C., Tisch, M., Metternich, J., Abele, E., 2017. Das Darmstädter Shopfloor Management-Modell. Zeitschrift für wirtschaftlichen Fabrikbetrieb 112 (3), 118–121.

[6] Peters, R., 2009. Shopfloor Management: Führen am Ort der Wertschöpfung. LOG_X, Stuttgart, 141 pp.

[7] Suzaki, K., 2014. New shop floor management: Empowering people for continuous improvement. Free Press.

[8] Hellebrandt, T., Ruessmann, M., Heine, I., Schmitt, R.H., 2019. Conceptual Approach to Integrated Human-Centered Performance Management on the Shop Floor, in: Kantola, J.I., Nazir, S., Barath, T. (Eds.), Advances in Human Factors, Business Management and Society, vol. 783. Springer International Publishing, Cham, pp. 309–321.

[9] Kantola, J.I., Nazir, S., Barath, T., 2019. Advances in Human Factors, Business Management and Society. Springer International Publishing, Cham, 710 pp.

[10] Ohlig, J., Hellebrandt, T., Poetters, P., Heine, I., Schmitt, R.H., Leyendecker, B., 2021. Human-centered performance management in manual assembly. Procedia CIRP 97, 418–422.

[11] Burdensky, D.M.A., Kneissl, B., Alt, R., 2018. Deskriptive Analyse von Kennzahlenrelationen: BMW Group, Leipzig München, Deutschland. Multikonferenz Wirtschaftsinformatik, 59–70.

[12] Kleindienst, B., 2017. Performance Measurement und Management: Gestaltung und Einführung von Kennzahlen- und Steuerungssystemen, 1. Auflage 2017 ed. Springer Fachmedien Wiesbaden, Wiesbaden, Online-Ressourcen.

[13] Diez, J.V., Ordieres-Mere, J., Nuber, G., 2015. The HOSHIN KANRI TREE. Cross-plant Lean Shopfloor Management. Procedia CIRP 32, 150–155.

[14] Lynch, R.L., Cross, K.F., 2003. Measure up!: How to Measure Corporate Performance, 2. ed., reprinted, transferred to digital print ed. Blackwell Business, Malden, Mass., 250 pp.
[15] Jourdan, N., Longard, L., Biegel, T., Metternich, J., 2021. Machine Learning For Intelligent Maintenance And Quality Control: A Review Of Existing Datasets And Corresponding Use Cases. Proceedings of the Conference on Production Systems and Logistics: CPSL 2021, 499–513.
[16] Meißner, A., Grunert, F., Metternich, J., 2020. Digital shop floor management: A target state. Procedia CIRP 93, 311–315.
[17] Longard, L., Meissner, A., Müller, M., Metternich, J., 2020. Digitales Shopfloor Management - Wohin geht die Reise? Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (9), 645–649.
[18] Wuest, T., Weimer, D., Irgens, C., Thoben, K.-D., 2016. Machine learning in manufacturing: advantages, challenges, and applications. Production & Manufacturing Research 4 (1), 23–45.
[19] Meister, M., Böing², T., Batz, S., Metternich, J., 2018. Problem-solving process design in production: Current progress and action required. Procedia CIRP 78, 376–381.
[20] Liker, J.K., Meier, D., 2006. The Toyota way fieldbook: A practical guide for implementing Toyota's 4Ps. McGraw-Hill, New York, 475 pp.
[21] Klamma, R., 2000. Vernetztes Verbesserungsmanagement mit einem Unternehmensgedächtnis-Repository. Technische Hochschule Aachen, Dissertation, 281 pp.
[22] Jung, B., Schweißer, S., Wappis, J., 2017. 8D - Systematisch Probleme lösen. Carl Hanser Verlag GmbH & Co. KG, München, 123 pp.
[23] Meissner, A., Müller, M., Hermann, A., Metternich, J., 2018. Digitalization as a catalyst for lean production: A learning factory approach for digital shop floor management. Procedia Manufacturing 23, 81–86.
[24] Bramer, M., 2020. Principles of Data Mining, 4th ed. 2020 ed. Springer London; Imprint Springer, London, 571 pp.
[25] Longard, L., Schiborr, L., Metternich, J., 2022. Potentials and obstacles of the use of data mining in problem-solving processes. Manuscript submitted for publication.
[26] Hackathorn, R., 2002. Minimizing Action Distance. DM Review, 1–5.
[27] Lanza, G., Hofmann, C., Stricker, N., Biehl, E., Braun, Y., 2018. Auf dem Weg zum digitalen Shopfloor Management: Eine Studie zum Stand der Echtzeitentscheidungsfähigkeit und des Industrie 4.0-Reifegrads.
[28] Ohlig, J., Hellebrandt, T., Metzmacher, A.I., Pötters, P., Heine, I., Schmitt, R.H., Leyendecker, B., 2020. Performance management on the shop floor – an investigation of KPI perception among managers and employees. International Journal of Quality and Service Sciences 12 (4), 461–473.
[29] Aggarwal, C.C., 2017. Outlier Analysis, 2nd ed. 2017 ed. Springer, Cham, 466 pp.
[30] Bishop, C.M., 2006. Pattern recognition and machine learning. Springer Science+Business Media LLC, New York, NY, 758 pp.
[31] Chang, I., Tiao, G.C., Chen, C., 1988. Estimation of Time Series Parameters in the Presence of Outliers. Technometrics 30 (2), 193–204.
[32] Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M., Sabeti, P.C., 2011. Detecting novel associations in large data sets. Science (New York, N.Y.) 334 (6062), 1518–1524.
[33] Peral, J., Maté, A., Marco, M., 2017. Application of Data Mining techniques to identify relevant Key Performance Indicators. Computer Standards & Interfaces 50, 55–64.

Biography

Lukas Longard, M. Sc. (*1992) is chief engineer and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2019. His research interests include the further development and introduction of machine learning to digital shop floor management and performance management systems.

Sebastian Bardy, M. Sc. (*1991) is PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2019. His research interests include the further development and introduction of resource-efficiency in digital shop floor management.

Prof. Dr.-Ing. Joachim Metternich (*1968) has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2012. He is spokesman of the medium-sized businesses digital competence centre Darmstadt, he has been president of the International Association of Learning Factories (IALF) from 2015 to 2021 and he is a member of the scientific society for production technology (WGP).

3rd Conference on Production Systems and Logistics

A Systematic Literature Review Of Machine Learning Approaches For The Prediction Of Delivery Dates

Janine Tatjana Maier[1], Alexander Rokoss[1], Thorben Green[1], Nikola Brkovic[1], Matthias Schmidt[1]

[1]*Institute of Product and Process Innovation, Leuphana University of Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany*

Abstract

Manufacturing companies tend to use standardized delivery times. The actual delivery times requested by the customers and the current capacity utilization of the production are often not taken into account. Therefore, such a simplification likely results in a reduction of the efficiency of the production. For example, it can lead to an obligation to use rush orders, an unrealistic calculation of inventories or an unnecessary exclusion of a Make-to-Order production. In the worst case, this results not only in an economically inadequate production, but also in a low achievement of logistic objectives and therefore in customer complaints. To avoid this, the delivery dates proposed to the customer must be realistic. Given the large number of customer orders, a wide range of products, varying order quantities and times, as well as various delivery times requested by customers, it is not economical to determine individual delivery dates manually. The ongoing digitalization and technological innovations offer new opportunities to support this task. In the literature, various approaches using machine learning methods for specific production planning and control tasks exist. As these methods are in general applicable for different tasks involving predictions, they can also assist during the determination of delivery dates. Therefore, this paper provides a comprehensive review of the state of the art regarding the use of machine learning approaches for the prediction of delivery dates. To identify research gaps the analyzed publications were differentiated according to several criteria, such as the overall objective and the applied methods. The majority of scientific publications addresses delivery dates only as a subordinate aspect while focusing on production planning and control tasks. Therefore, the interrelationships with several production planning and control tasks were considered during the analysis.

Keywords

Delivery Date; Production Planning And Control; Prediction; Machine Learning; Literature Review

1. Introduction

The level of customer expectations regarding the logistic performance has strongly increased in the last decades. Nowadays, customers expect not only individual products with a high quality for low prices, but also short and especially reliable delivery times [1]. This results in major challenges for manufacturing companies. From an economical point of view, operating a finished goods store might not be beneficial due to the large number of products and product variants. Therefore, companies tend to use standard delivery times within an Assemble-to-Order or a Make-to-Order production [2]. This leads to a strong fluctuation of the capacity utilization and requires a high amount of effort for continuously adjusting the capacity. In addition, the requirements of customers can vary greatly. While in some cases the delivery should just be as fast as possible, others demand a specific time window for the delivery or even a just-in-time delivery.

DOI: https://doi.org/10.15488/12157

Customers tend to buy more and in some cases are willing to pay a significantly higher price in case of short and reliable delivery times [3]. Therefore, in practice concepts such as rush orders or delivery time classes with corresponding price differences exist. In the literature, the focus lies on the design of the products and production processes. Even though many approaches such as product platforms exist [4], authors also highlight the importance of the negotiation process in supply chains [5]. As the price and the delivery date are stated to be the two most critical factors in various industries [6], the assignment of delivery dates has been addressed by some authors [7]. However, the majority of these approaches is based on numerous assumptions and has been published several decades ago. Therefore, they do not reflect the high dynamics of today's markets. To compete successfully, manufacturing companies need to predict their delivery dates continuously, quickly, and realistically. In this context customer heterogeneity provides a challenge but also offers opportunities [8]. Customers expect a fast estimation of the delivery date, especially if the delivery time and the price per unit can be negotiated [9] or if they need to be present in person to receive the delivery [10]. Looking at the large amount of customer orders and products combined with varying order quantities and times, predicting delivery dates manually is impossible from an economic point of view. The need of assistance can also be seen in the development of numerous decision support systems [11].

Machine learning (ML) methods provide the possibility to process the necessary high amount of data and are generally suitable for applications involving predictions [12]. As data availability, consistency and integrity increase, the use of these methods to solve production planning and control (PPC) tasks is becoming increasingly popular [13] [14]. The aim is commonly to optimize a system with predefined delivery dates. This includes tasks like the order acceptance [15], order release [16], or sequencing [17]. All these tasks relate at least indirectly to the prediction of delivery dates. The throughput time is not only a main part of the delivery time, it also interacts with upstream decisions such as the selection of the order processing strategy [18]. Although many researchers have already addressed this topic and the use of ML in PPC in general, the need for research in this area remains. A recent study shows that about 75% of the possible research domains for ML in PPC have been examined only to a minor extent or not at all [19].

In this paper, a systematic literature review of ML approaches for the prediction of delivery dates is presented. The subsequent section outlines the applied procedure. This includes the research questions, the selection criteria, the quality assessment and the data analysis of the systematic literature review. The selected publications were differentiated and examined according to several criteria. The results of the analysis are presented and discussed in section three. Lastly, in section four a summary is given and research gaps are highlighted.

2. Method and data

The study aims for a comprehensive overview of the state of the art regarding the prediction of delivery dates and the identification of research gaps. A systematic literature research was conducted following the guidelines of Kitchenham [20]. In the following, the individual steps are outlined.

2.1 Research questions

The following research questions (RQ) were raised:

- RQ1: What research topics are being addressed?
- RQ2: Which use cases are being addressed?
- RQ3: Which methods have been used?
- RQ4: Do ML models outperform non-ML models?
- RQ5: Which trends are recognizable?
- RQ6: What are the limitations of current research?

RQ1 aims to identify interrelationships between the prediction of delivery dates and PPC tasks as well as upstream decisions and general aspects of production management. RQ2 focuses on how suitable the theoretical research results are for industrial practice and to what extent they have already been implemented. Answering RQ3 and RQ4 provides insights on the methods and helps to identify research gaps. Based on the results gained through RQ1 to RQ4, specific aspects are selected for a detailed analysis. To pinpoint the trends mentioned in RQ5, the timeline is examined with regards to certain innovations and changes, such as the introduction of the term industry 4.0 in 2011. RQ6 is directly related to all other research questions and thus calls for a critical review of the previous results.

2.2 Search process and selection criteria

The search was carried out using the databases Scopus and Web of Science. As these databases are known for their scientific relevance, they are widely used for literature reviews [21]. Regarding the disciplines of economics and engineering, they have numerous overlaps, but are not completely identical [22]. Figure 1 gives an overview of the selection process for the systematic literature review containing the selection criteria as well as the results returned from the databases.

Step	Limitating aspects	Criteria	Results:	
1	Titel, abstract, key words + Publication year	„Delivery date“ or synonyms from 2002 to 2021	Scopus: 29,629	Web of Science: 12,813
2	Titel, abstract, key words	„Machine learning“ or synonyms	Scopus: 1,705	Web of Science: 601
3	Source	Journals or conference proceedings	Scopus: 1,625	Web of Science: 598
4	Language	Written in English	Scopus: 1,591	Web of Science: 593
5	Research area	Related to production management	Scopus: 810	Web of Science: 360
6	Content alignment	Title, abstract	Scopus: 100	Web of Science: 89
7	Content alignment + Quality assessment	Full paper, No duplicates, full text available, Quality assessment questions	Total: 62	

Figure 1: Selection process during the systematic literature review

The initial step was to search the term “delivery date” and its synonyms in the title, abstracts and keywords of the years 2002 to 2021. As stated above the delivery date and the throughput time are strongly related. Therefore, besides “due date” and “delivery time”, the terms “throughput time” and “lead time” were also considered to be synonyms. PPC tasks not directly related to the prediction of delivery dates, such as order acceptance, order release and scheduling, were not considered at this point as their interrelations were taken into account during the examination of the selected papers. A previously conducted study revealed a noticeable growth in scientific publications regarding the use of ML methods in PPC starting in 2007 [23]. To ensure the identification and evaluation of trends and at the same time enable a detailed and efficient analysis of the current state of the research on the prediction of delivery dates the final time horizon was set to be 20 years. As the term PPC as well as its current understanding were established around 40 years ago, e.g. by the PPC model of Hackstein in 1984 [24], analyzing a longer period does not seem to be suitable. This is strengthened by authors suggesting the use of intelligent systems for planning production processes, such as Mill and Spraggett in 1984 [25] or Yang et al. in 1992 [26].

In the basic literature at this time the determination of due dates was focused on internal due dates as part of production scheduling [7] [27]. The delivery date was classified as an external factor as it is decided by the customer or by the sales department.

The next step was to link the terms referring to prediction using ML methods. The Boolean AND as well as the Boolean OR were used to incorporate synonyms and alternative spellings. This resulted in the terms "machine learning", "deep learning", "neural network", "artificial intelligence", "data analytics" and "data mining". Although these terms have different meanings, they are often used synonymously in practice [28]. To obtain high-quality publications and at the same time avoid the repetition of content the sources should be limited. It is common to select only articles from scholarly journals [29] and conference proceedings [30] as the majority of these have been peer-reviewed prior to publication [31]. To provide the basis for a detailed evaluation of the full papers the results were limited to papers written in English. Afterwards, all topics not related to production management were excluded as they may use the same terms in a different context. This was followed by the evaluation of the titles and abstracts regarding the content alignment. The databases were compared and duplicates as well as papers with no full text available were removed. Lastly, the full papers were evaluated regarding the content alignment and the quality assessment. Papers with a quality score of less than 5 were excluded from the study.

2.3 Quality assessment and data analysis

To ensure a high quality the relevance, credibility and rigorousness of the selected studies need to be checked [32]. The following ten quality assessment questions (QAQ) were applied [32] [33]:

- QAQ1: Does the study report empirical research or is it a report based on the opinion of an expert?
- QAQ2: Are the aims and the motivation of the research clearly defined?
- QAQ3: Is the estimation context adequately described?
- QAQ4: Are the methods well defined and deliberate?
- QAQ5: Is the research design appropriate and justifiable?
- QAQ6: Does the study contain a sufficient project data set?
- QAQ7: Is the proposed method compared to other methods?
- QAQ8: Are the findings of the study clearly stated and supported by reporting results?
- QAQ9: Are the limitations of the study analyzed explicitly?
- QAQ10: Does the study provide value for academia or industrial practice?

The answers were scored as "No" = 0, "Partial" = 0.5 and "Yes" = 1. For each selected paper, data regarding the topic, e.g. title, key words, main area and related topics mentioned, the authors, the source, the study type, the methods used as well as the quality evaluation were extracted.

3. Results

The literature research resulted in 62 papers addressing the use of ML methods in the context of the prediction of delivery dates. 33 (53%) papers appeared in scientific journals, while 29 papers (47%) were published in conference proceedings. The papers were classified based on the five main categories:

- negation processes
- time periods
- methods
- data
- PPC tasks

Each paper could be assigned to several main and sub categories.

3.1 Topics and methods

As high logistic performance is a relevant purchasing criteria for today's customers, the adherence to delivery dates is a highly discussed topic in the literature. The complexity of related topics such as scheduling and routing is enhanced by the uncertainties in sales forecasting, production problems and delays in delivery existing in industrial practice. The delivery date is strongly influenced by these uncertainties. Nevertheless, the delivery date prediction is mainly considered to be a subordinate aspect of the logistic performance. Therefore, papers calculating due dates while mainly focusing on the optimization of a system, e.g. order release, scheduling or inventory management to minimize costs, were excluded from this study. In case the determination of specific time periods of orders was conducted to negotiate the delivery date with the customer the papers were considered relevant for the topic. Figure 2 shows the distribution of the papers regarding addressed time periods and PPC tasks (RQ1).

Sorting the papers by the time period addressed revealed that the delivery time is mainly an important feature regarding last mile delivery, such as package delivery or shipment processes, and the negotiation processes between manufacturers, suppliers and customers. The prediction of the throughput time and its components processing time and inter-operation time are the main topic focused within the context of delivery dates. In addition, a few authors highlight unexpected delays, for example due to machine breakdowns and the related determination of safety times. The highly varying interest in time periods also reflects in considered interrelationships with related PPC tasks. The acceptance or rejection of an order depends on the negotiation process and its features such as the price per unit und the delivery time. Therefore the amount of papers dealing with the order acceptance is similar to the ones focusing directly on the delivery time. In a standard Make-to-Order production the lot size is equal to the size of the customer order and there is no semi-finished or finished goods store. Therefore, the PPC tasks lot sizing and inventory planning are not directly relevant for the determination of the delivery time. Nevertheless, they are a few times addressed in the context of the dispatch time as well as an influencing factor during scheduling. As the throughput time and its components are the mostly investigated topic, the directly related PPC tasks scheduling and capacity planning occur in various studies.

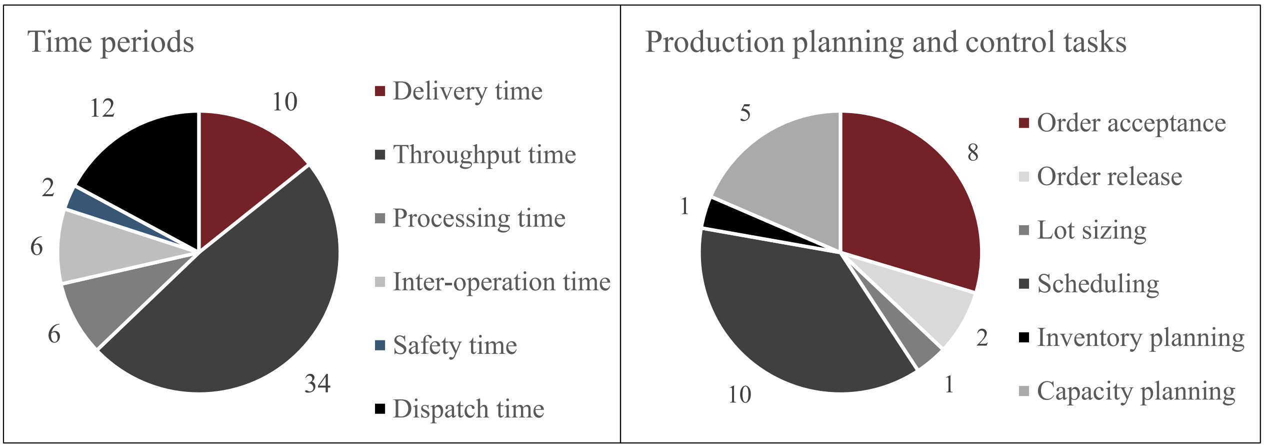


Figure 2: Papers assigned by the addressed time period and the related production planning and control tasks

The applied methods are as broad as the addressed time periods and PPC tasks (RQ3). They range from fundamental mathematical models to concepts based on a combination of different ML methods. Neural networks are the most used method as various versions of them appear in 45% of the papers. However, this high percentage can be explained by the fact that most of the authors present a comparison of several methods for a specific problem (79%) rather than a new universal approach (21%). In cases where ML methods were compared with conventional methods, they generally performed better (RQ4). However, most of the approaches considered only a few or even just one objective.

This can also be seen in the data sets (RQ2). Only one paper described a concept without proving it by a numerical example. About a third of the authors referred to simple virtual data sets for their proof of concept, while the rest used a case study showing the applicability of their approach in industrial practice (65%). Various industries are covered through logistic companies and suppliers, typical manufacturing companies, such as automotive manufacturers, shipyards or semiconductor manufacturers, and e-commerce platforms.

3.2 Trends

Figure 3 shows the annual number of publications for the years 2002 to 2021 using a bar chart (RQ5). Although some variations are evident, there is an overall increase in the number of publications during the examined period. Starting from 2014 a continuous growth is visible. This could be related to the introduction of the term “Industry 4.0” at the Hannover Fair in 2011 and with the increased use of ML methods such as "deep learning".

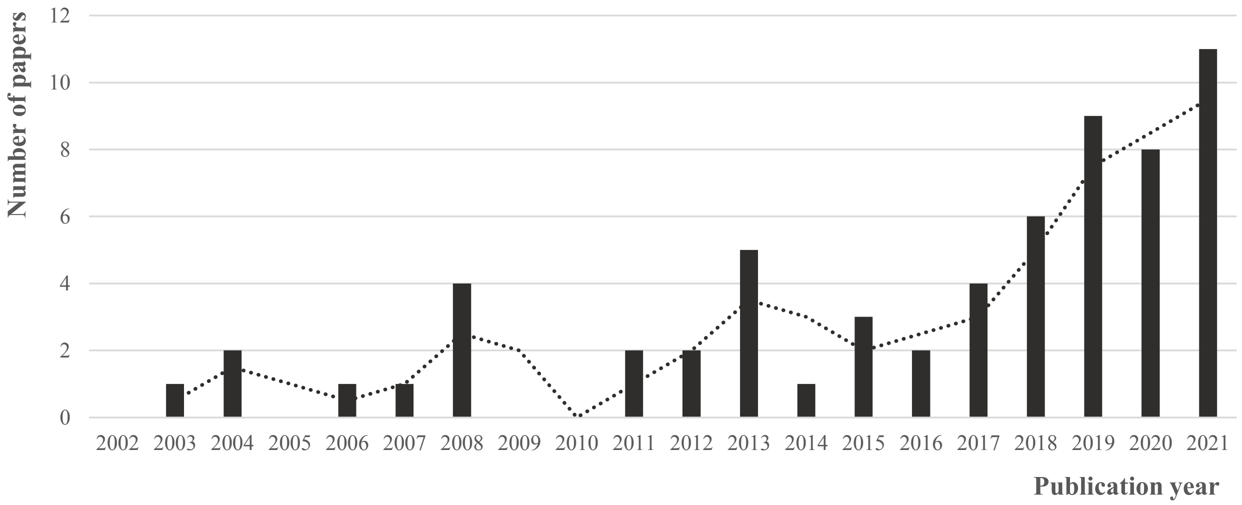


Figure 3: Distribution of papers from 2002 to 2021

The examination of the geographical distribution revealed that publications originated from a total of 24 different countries. The five countries with the highest amount of papers are Taiwan (11 papers), China (10 papers), Germany (9 papers), the United States (8 papers) and Austria (7 papers). This slight imbalance can be explained by authors presenting extensions of their own approaches and using the same data sets.

The keywords were analyzed using the software VOSviewer [34]. The so-called visualization of similarities (VOS) maps can be used to represent relationships between objects in various ways. The strength of the linkage determines the location of the keywords within the VOS maps. The size of the points assigned to the keywords correlates with the number of occurrences of the respective keyword. A keyword was considered relevant to the topic if it appeared at least two times. This assumption resulted in one group containing 133 connected keywords (Figure 4). The multiple cross-linking of the individual keywords highlights the strong connection between PPC, ML methods and the delivery time.

Arranging the keywords by year, reveals a minor change in the terms over the time. Mathematical models and decision-making based on conventional rules tends to be replaced by ML methods. The focus seems to start shifting from the throughput time to a more universal view including smaller time periods like transitions times and travel times. As the increasing customer requirements regarding the logistic performance require the accurate prediction of delivery dates, forecasting delivery times draws attention towards the handling and the quality of data.

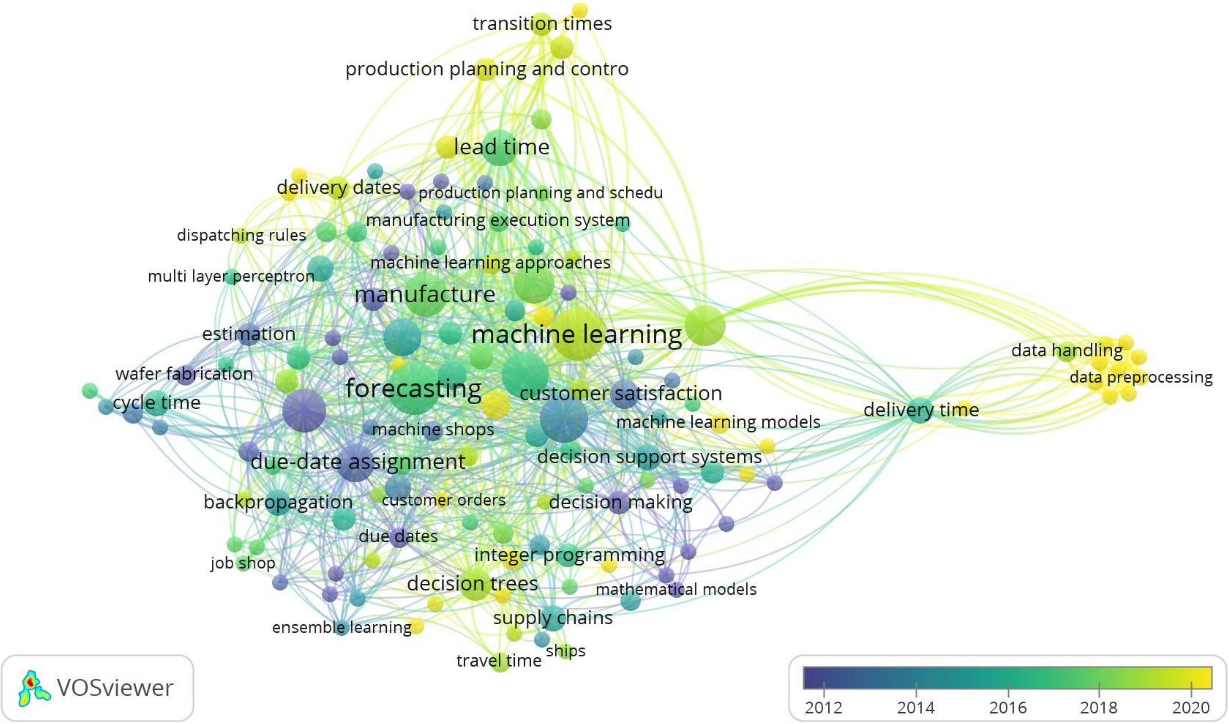


Figure 4: Results of the keyword analysis using VOSviewer

4. Conclusions and outlook

The increasing number of publications related to the prediction of delivery dates with ML methods in the last 20 years reflects the importance of this topic and the growing interest of academia and industry in it. In terms of content, the publications primarily focus on the throughput time. Nevertheless, a reorientation towards previously neglected time components of the delivery time is recognizable. This can be explained by the large number of already existing approaches concerning the determination of the throughput time and its optimization by scheduling as well as the progress made in the area of ML, and thus the simplified application to more complex problems.

In summary, a strong interrelation between the determination of delivery dates, ML methods and PPC tasks is visible. There are various fields of application and the number of publications in this area will probably keep increasing in the next years as there is still a strong imbalance leaving a research gap. Detailed analysis of the different time periods related to the delivery time as well as a holistic model for the prediction of delivery dates is required. As an initial step towards such a model, various case studies are required. There is ongoing research with partners from industrial practice on process quality, pricing, sales planning and storage dimensioning. In addition, examining the various existing interrelations with PPC tasks as well as upstream strategic decisions such as the selection of the order processing strategy or the location of production sites and warehouses could provide interesting insights. To benefit the prediction of delivery dates appropriately additional research is required in the area of forecasting customer demand and behavior as well as regarding the options offered to customers upfront like rush orders.

Acknowledgements

The study was carried out in the framework of the research project „KI-Werkstatt Künstliche - Intelligenz in Produktionsunternehmen (KIWe)“. It was supported by the European Union and the State of Lower Saxony within the EFRE program.

References

[1] Paprocka, I., Cyba, S., 2015. Assessment of Production Capacity and Ability of Rapid Response to Changing Customer Expectations. Applied Mechanics and Materials 809, 1378–1383.

[2] Rao, U.S., Swaminathan, J.M., Zhang, J., 2005. Demand and production management with uniform guaranteed lead time. Production and Operations Management 14 (4), 400–412.

[3] Rao, S., Griffis, S.E., Goldsby, T.J., 2011. Failure to deliver? Linking online order fulfillment glitches with future purchase behavior. Journal of Operations Management 29 (7-8), 692–703.

[4] Simpson, T. W., 2004. Product platform design and customization: Status and promise. Artificial Intelligence for Engineering Design, Analysis and Manufacturing 18 (1), 3–20.

[5] Pan, A., Choi, T.-M., 2016. An agent-based negotiation model on price and delivery date in a fashion supply chain. Annals of Operations Research 242, 529–557.

[6] Peng, D. X., Lu, G., 2017. Exploring the impact of delivery performance on customer transaction volume and unit price: evidence from an assembly manufacturing supply chain. Production and Operations Management 26 (5), 880–902.

[7] Cheng, T.C.E., Gupta, M.C., 1989. Survey of scheduling research involving due date determination decisions. European Journal of Operational Research 38, 156–166.

[8] Wijekoon, A., Salunke, S., Athaide, G.A., 2021. Customer heterogeneity and innovation-based competitive strategy: A review, synthesis, and research agenda. Journal of Product Innovation Management 38 (3), 315–333.

[9] Rau, H., Tsai, M.H., Chen, C.W., Shiang, W.J., 2006. Learning-based automated negotiation between shipper and forwarder. Computers & industrial engineering 51 (3), 464–481.

[10] Cwioro, G., Hungerländer, P., Maier, K., Pöcher, J., Truden, C., 2019. An optimization approach to the ordering phase of an attended home delivery service, in: Rousseau, L.M., Stergiou. K. (Eds.) Integration of Constraint Programming, Artificial Intelligence, and Operations Research. CPAIOR 2019. Lecture Notes in Computer Science 11494. Springer, Cham, pp. 208–224.

[11] Sala, R., Pezzotta, G., Pirola, F., Huang, G.Q., 2019. Decision-support system-based service delivery in the product-service system context: Literature review and gap analysis. Procedia CIRP 83, 126–131.

[12] Mayr, A., Kißkalt, D., Meiners, M., Lutz, B., Schäfer, F., Seidel, R., Selmeier, A., Fuchs, J., Metzner, M., Blank, A., Franke, J. (2019). Machine Learning in Production–Potentials, challenges and exemplary applications. Procedia CIRP 86, 49–54.

[13] Reuter, C., Brambring, F., Weirich, J., Kleines, A., 2016. Improving data consistency in production control by adaptation of data mining algorithms. Procedia CIRP 56, 545–550.

[14] Schuh, G., Reuter, C., Prote, J.P., Brambring, F., Ays, J., 2017. Increasing data integrity for improving decision making in production planning and control. CIRP Annals–Manufacturing Technology 66 (1), 425–428.

[15] Zhang, H., Leng, J., Zhang, H., Ruan, G., Zhou, M., Zhang, Y., 2021. A deep reinforcement learning algorithm for order acceptance decision of individualized product assembling. 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), 21–24.

[16] Schneckenreither, M., Haeussler, S., Gerhold, C., 2021. Order release planning with predictive lead times: a machine learning approach. International Journal of Production Research 59 (11), 3285–3303.

[17] Liang, Y.C., Lee, Z.H., Chen, Y.S., 2012. A novel ant colony optimization approach for on-line scheduling and due date determination. Journal of Heuristics 18(4), 571–591.

[18] Stevenson, M., Hendry, L.C., Kingsman, B.G., 2005. A review of production planning and control: the applicability of key concepts to the make-to-order industry. International Journal of Production Research 43(5), 869–898.

[19] Cadavid, J.P.U., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A., 2020. Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. Journal of Intelligent Manufacturing 31, 1531–1558.

[20] Kitchenham, B., 2004. Procedures for Undertaking Systematic Reviews, Joint Technical Report, Computer Science Department, Keele University (TR/SE0401) and National ICT Australia Ltd. (0400011T.1)

[21] Mongeon, P., Paul-Hus, A., 2016. The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics 106 (1), 213–228.

[22] Martín-Martín, A., Orduna-Malea, E., Thelwall, M., López-Cózar, E.D., 2018. Google Scholar, Web of Science, and Scopus. A systematic comparison of citations in 252 subject categories. Journal of Informetrics 12 (4), 1160–1177.

[23] Schmidt, M., Maier, J.T., Grothkopp, M., 2020. Eine bibliometrische Analyse: Produktionsplanung und -steuerung und maschinelles Lernen. wt Werkstattstechnik online 110 (4), 220–225.

[24] Hackstein, R., 1984. Produktionsplanung und -steuerung (PPS). Ein Handbuch für die Betriebspraxis. Düsseldorf: VDI-Verlag

[25] Mill, F., Spraggett, S.,1984. Artificial intelligence for production planning. Computer-Aided Engineering Journal 1(7), 210–213.

[26] Yang, H., Lu, W.F., Lin, A.C., 1992. Intelligent Process Planning Using a Machine Learning Approach. IFAC Proceedings 25 (28), 147–151.

[27] Karmaker, U.S., 1987. Lot Sizes, Lead Times and In-Process Inventories. Management Science 33(3), 409–418.

[28] Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science 349 (6245), 255–260.

[29] Rowley, J., Slack, F., 2004. Conducting a literature review. Management Research News 27 (6), 31–39.

[30] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare For the Future: Writing a Literature Review. MIS Quarterly 26 (2), xiii-xxiii

[31] vom Brocke, J., Simons, A., Niehaves, B., Plattfaut, R., Cleven, A., 2009. Reconstructing the giant: on the importance of rigour in documenting the literature search process. ECIS 17th European Conference on Information Systems, 2–13.

[32] Dybå, T., Dingsøyr, T., 2008. Empirical studies of agile software development: A systematic review. Information and software technology 50 (9–10), 833–859.

[33] Wen, J., Li, S., Lin, Z., Hu, Y., Huang, C., 2012. Systematic literature review of machine learning based software development effort estimation models. Information and Software Technology 54 (1), 41–59.

[34] Van Eck, N., Waltman, L, 2009. Software survey. VOSviewer, a computer program for bibliometric mapping. Scientometrics 84 (2), 523–538.

Biography

Janine Tatjana Maier (*1994) studied industrial engineering at the Leibniz University Hannover. Since 2018, she works as a research associate in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University of Lüneburg.

Alexander Rokoss (*1993) is a research associate and PhD-student in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University Lüneburg since 2018. His research focus is on the practical application of machine learning methods in production planning and control.

Thorben Green (*1991) studied industrial engineering at the Leuphana University of Lüneburg. Since 2021, he works as a research associate in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University of Lüneburg.

Nikola Brkovic (*1994) studies Management & Engineering at the Leuphana University Lüneburg. Since 2020, he supports the chair of production management as a student research assistant in various projects at the Institute of Product and Process Innovation (PPI).

Matthias Schmidt (*1978) studied industrial engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he became head of Research and Industry of the IFA and received his habilitation. Since 2018, he holds the chair of production management at the Institute for Product and Process Innovation (PPI) at the Leuphana University of Lüneburg. In addition, he became the head of the PPI in 2019.

3rd Conference on Production Systems and Logistics

Tool Wear Prediction Upgrade Kit For Legacy CNC Milling Machines In The Shop Floor

Yuechi Jiang[1], Benny Drescher[1,2], Max Wittstamm[1,3], Cuihong Hu[1], Florian Clemens[4], Weimin Wang[1], Volker Stich[4]

[1]Hong Kong Industrial Artificial Intelligence and Robotics Centre (FLAIR), Hong Kong, SAR China
[2]Invention Center at RWTH Aachen Campus (INC), Aachen, Germany
[3]Center Connected Industry at RWTH Aachen Campus (CCI), Aachen, Germany
[4]FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany

Abstract

The operation of CNC milling is expensive because of the cost-intensive use of cutting tools. The wear and tear of CNC tools influence the tool lifetime. Today's machines are not capable of accurately estimating the tool abrasion during the machining process. Therefore, manufacturers rely on reactive maintenance, a tool change after breakage, or a preventive maintenance approach, a tool change according to predefined tool specifications. In either case, maintenance costs are high due to a loss of machine utilization or premature tool change. To find the optimal point of tool change, it is necessary to monitor CNC process parameters during machining and use advanced data analytics to predict the tool abrasion. However, data science expertise is limited in small-medium sized manufacturing companies. The long operating life of machines often does not justify investments in new machines before the end of operating life. The publication describes a cost-efficient approach to upgrade legacy CNC machines with a Tool Wear Prediction Upgrade Kit. A practical solution is presented with a holistic hardware/software setup, including edge device, and multiple sensors. The prediction of tool wear is based on machine learning. The user interface visualizes the machine condition for the maintenance personnel in the shop floor. The approach is conceptualized and discussed based on industry requirements. Future work is outlined.

Keywords

CNC milling; predictive maintenance; condition monitoring; Tool Condition Monitoring (TCM); tool wear prediction; industry 4.0

1. Introduction

Milling is one of the most widespread manufacturing processes in the industry. A spindle motor rotates a multitooth tool to produce a variety of workpiece surfaces through Computerized Numerical Control (CNC) movements [1,2]. Forces and friction at the cutting tool are caused by high-power machining and the process-related interruption of the cut for each cutting edge [1]. To control the overall product quality and operate at high performance, it is necessary to change tools frequently [3]. Tool maintenance accounts for 20% of machine downtime and 25% of total machining costs [4]. An unexpected one-day stoppage of a milling machine caused by a tool breakage could lead to costs of about 100,000 – 200,000 EUR [5]. Due to the necessity to operate at maximum yield, finding the optimal point of time for tool change is technically challenging [5], but critical for Overall Equipment Effectiveness. On one side, premature tool change leads

DOI: https://doi.org/10.15488/12150

to efficiency losses. On the other side, tool wear leads to product quality loss, tool breakage and potential machine stoppage [6]. Tool condition monitoring promises to provide the necessary maintenance information by diagnosing the tool condition in real-time and thereby enabling the prediction and scheduling of tool changes [7,3]. In contrast to time-based maintenance strategies, the assessment is based on real-time collected sensor data in the process, instead of rigid intervals from tool suppliers or experience from maintenance technicians or operators [5]. Condition-based maintenance has been a topic in research for many years, but practical solutions are limited, as most approaches are theoretical, and implementations are conducted in laboratory environments [2]. Small and medium-sized enterprises (SME) especially struggle with the applicability because they often have deficits of experience, competence, and capital [8]. Commercially available off-the-shelve condition monitoring solutions are technically limited due to integration barriers, proprietary interfaces, software performance, flexibility issues, and high capital costs [2]. The machine park of SMEs is characterized by a variety of machines from different vendors which do not have appropriate sensing technologies for tool wear prediction and network connectivity available [9,10]. Required expertise in data science is rare [11]. Hence, practical approaches need to be developed that can be installed at various legacy machines. They shall be integrated into shop floor IT infrastructures and have an intuitive user interface. This research work conceptualizes a Tool Wear Prediction Upgrade Kit that can be applied to legacy CNC milling machines. Related work in tool wear measurement systems, tool wear prediction architectures, and Machine Learning (ML) is outlined (section 2). The system architecture and application methodology are described (section 3). The concept is discussed (section 4) and future work is concluded (section 5).

2. Related work

2.1 Tool wear measurement systems

Tool wear occurs at the flank and chip side of cutting tools. On the flank side, friction between the tool and the surface of the workpiece results in tool flank wear. On the chip side, crater wear is caused by the movement of the chip on the tool. Among the two types, flank wear is considered the predominant evaluation index because of the effects on the workpiece quality and process reliability. [12] Tool wear can be measured using direct and indirect methods. Direct methods measure the tool wear using imaging devices, such as a microscope or a machine vision system [2]. This achieves high accuracy but causes machine downtime due to the measurement [13]. Indirect methods rely on sensor data of machining processes to estimate tool wear conditions. Signals relevant to tool wear in milling processes include cutting forces, acoustic emissions, vibrations, power consumption, temperature, and sounds [2]. A dynamometer can measure the tangential and axial cutting forces applied to the workpiece, which is highly relevant to tool wear, as the decrease of tool sharpness causes the increase of cutting forces [13]. However, the dynamometer has direct contact with the workpiece. Electric current sensors can be an alternative to dynamometers, as the increase of cutting forces causes an increase of motor current [14]. Acoustic emission sensors measure the strain wave inside the workpiece. Recent studies show that it provides the most effective information for tool flank wear prediction [13]. Accelerometers can measure the vibration of the spindle and the workpiece, as the decrease of tool sharpness causes an increase of vibrational energy [15]. Temperature sensors measure the heat generated during milling. On the one side, the milling process generates high temperatures. On the other side, high temperatures cause tool wear. It is shown that temperature and tool wear have a positive correlation [16]. However, temperature measurement is affected by lubrication [17], and thus is ineffective for practical application during machining. Microphones measure the sound generated during milling. Studies show that an increase of tool wear causes an increase of sound intensity [18]. Sound signals may face interference from other machine noises; thus, they have limited applicability in noisy shop floors. Besides the relevance of sensor signals, the installation of sensors should also be considered. The sensors should be close to the milling surface without interfering with the milling process. Spindle housing and workpiece fixtures are feasible positions for sensor installation [2]. Direct measurements are time-consuming, often requiring an additional manual step and cannot run in parallel to machining processes [13]. Thus, the research work concentrates on indirect measurements instead of direct measurements. From literature, it can be derived that the indirect measurement of acoustic emissions, vibration sensors and electric motor current provide effective information for tool flank wear. The sensors can be mounted at CNC machines by a simple installation procedure and can be equipped on various machine types and vendors. Therefore, the Tool Wear Prediction Kit's sensor system consists of acoustic emission, vibration, and electric motor current. An increase in accuracy is expected due to the combination of multiple sensor types for the tool wear estimation.

2.2 IT architectures for tool wear prediction

Two approaches of system architectures for estimating the tool wear of CNC milling machines are identified. The first approach includes data pre-processing, feature engineering, modeling, and evaluation of the ML algorithms [19–22]. It considers the development of algorithms but excludes the data acquisition, storage, implementation, and productive operation. Thus, it is only applicable in corporate practice if the associated infrastructure for data acquisition, storage, and result visualization is already in place. The second approach additionally includes data acquisition, storage, and visualization of predicted results. Based on Rastegari et al. [23], vibration sensors are mounted on the CNC machine's spindle to measure its vibration when the spindle moves horizontally. The sensors are connected by cable to a measurement system unit and transformed into a digital signal. From the measurement system unit, the digitized sensor data is transferred to a database via 3G-Network. The stored data is collected and used for the training of the algorithms. Subsequently, the prediction results are visualized. Gouarir et al. [24] present an architecture based on an in-process prediction approach. A dynamometer is mounted to measure the cutting forces. The signal is amplified and sent to a data acquisition system via a cable. The data is stored in an experience database, which is used to train the prediction algorithms. The prediction results are then visualized in a human-machine interface. The architectures of the second approach are based on individual and experimental setups and do not include the productive operation and scaling of the algorithms in corporate practice. Thus, tool wear prediction is still hardly used in corporate practice, although researchers constantly improve the required ML algorithms. Therefore, this paper describes a scalable, holistic architecture that takes the productive operation for multiple machines in corporate practice into account. An edge device running relevant software is used to facilitate applicability. The software design supports the scalability by using lightweight IoT protocols, databases, and device management.

2.3 Machine learning algorithms for tool wear prediction

The ML algorithms need to consider two aspects in predicting tool wear. The first aspect is data pre-processing. Data pre-processing aims to extract representative features from the raw sensor data. Therefore, different sensor data may require different data pre-processing methods. The second aspect is modeling. Modeling aims to build a mathematical model to predict tool wear based on the input features. Thus, different input features may require differentmodeling methods. The recorded sensor data are time-series signals, and thus time-domain features can be extracted from the raw sensor data. Widely used time-domain features include the average value, the standard deviation value, the root mean square value, the kurtosis, and the skewness. For fast-fluctuating sensor data, such as vibration signals, there exists frequency information, and thus frequency-domain features can be extracted using Fourier transformation, such as the power spectrum and the spectral entropy. For non-stationary signals, the frequency spectrum may change over time. In this situation, it is suitable to extract time-frequency features from the raw sensor data, such as the wavelet transform features. [12] Besides extracting features from a single sensor, it is helpful to extract different features from multiple sensors [14]. Feature selection methods can be applied to keep the most useful features and remove useless features, such as principal component analysis, Pearson correlation analysis, and the monotonicity of the features [25,26]. After extracting representative features from the raw sensor data, a prediction model is trained to predict the tool wear. Widely used prediction models include support vector regression, fuzzy inference system, extreme learning machine, and artificial neural networks [13,27]. By extracting the frequency spectrum of the raw sensor data, convolutional neural networks can be applied, and multiple sensor data can be fused into different channels of the convolutional neural network [28]. In order to extract the sequential information from the sensor data, recurrent neural networks with the long short-term memory unit have been proven to outperform many traditional features for tool wear prediction [29]. In addition, 1D convolutional neural network can also extract sequential information [30].

The prediction models are data-driven models, and thus highly depend on the quality of the collected data. Simple models, such as linear regression and support vector regression require fewer data but have lower fitting capability. Complex models, such as neural networks, have higher fitting capability but require more data. Particularly, determining the optimized network structure, e.g., the number of hidden neurons and hidden layers in a neural network, is challenging. The interpretability of data-driven models still need improvement. Thus, it is promising to combine data-driven models with traditional physical models to improve the reliability. [11] A variety of ML models for tool wear prediction have been investigated in research. It can be observed that the selection of the most accurate model with the best performance depends on the conditions of the environment-specific milling process. There exists no one common analysis framework [12]. Therefore, a methodology to select and set the ML model for the application of the system is required.

2.4 Requirements

Related publications investigate tool wear monitoring systems only partially. Despite reasonable advancements in literature, state-of-the-art solutions mostly focus on experimental machine setups and lack an application in industry. A holistic industry-oriented system is yet missing. Therefore, a Tool Wear Prediction Upgrade Kit for legacy CNC milling machines in an industrial setting is developed within this publication. The research work has the following requirements:

a) **Independence of CNC machine vendor:**
The Upgrade Kit shall reach a high degree of sovereignty from CNC vendors; among others, it must avoid using closed protocols or non-standard interfaces to legacy machines.

b) **Adaptability to company-specific CNC machines:**
The Upgrade Kit shall be adjustable to the various CNC machine environments in the industry.

c) **Practical deployment and system scalability:**
The hardware and software framework shall be established as a platform and can be deployed to multiple machines with minor adjustments in system settings or configurations. Practical and easy deployment shall reduce setup time and underline a wide use of the Upgrade Kit.

d) **Stable and industrial-grade system design:**
The Upgrade Kit shall be capable of operating continuously 24 hours daily without supervision. The hardware follows industrial standards for deploying in a factory environment.

e) **Easy usage by maintenance team:**
The Upgrade Kit shall be used easily by the maintenance team in order to improve daily operations. Remote monitoring capabilities shall reduce visual check-ins at CNC machines or with operators on-site.

3. Tool Wear Prediction Upgrade Kit

3.1 System description

The architecture consists of the modules “User Interface”, “Software Hub” and “Machine Learning Model Tool Wear Prediction”, see Figure 1. The solution is executed on an edge device which enables low latency and fast data processing. The module “User Interface” provides components to depict live data and results from the tool wear prediction. The module “Software Hub” provides data collection, storage, processing, and device management components. The module “Machine Learning Model Tool Wear Prediction” can be divided into the productive use of the model and the training of the model. It provides components to pre-process the data, train algorithms and implement them for productive operation.

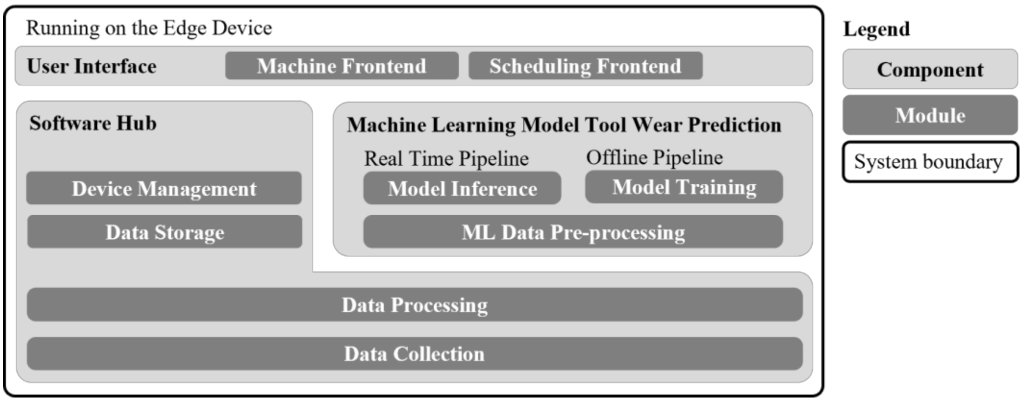


Figure 1: System architecture

As shown in Figure 2, an acoustic emission sensor, a vibration sensor, and a current sensor are used to enable data collection. Acoustic emission and vibration are measured for the spindle of a CNC machine. To process the sensor's analogue signal, a data acquisition system (DAQ) is connected to the sensors. After the transformation of the analog signal, the data is sent to the edge device for temporary storage and potential pre-processing. Raw data or pre-processed data are further transmitted to a central server via Message Queuing Telemetry Transport protocol (MQTT). The communication is organized by an MQTT broker. Utilizing MQTT ensures flexibility, modularity, and ease of implementation. [31] To realize an efficient communication, Sensor Markup Language (SenML) data format is used [32]. The edge device is used for providing computational capabilities for pre-processing as well as model training and data storage. Therefore, the edge device provides an InfluxDB time-series database as well as a PostgreSQL relational database. The time-series database is utilized to store the measurements and tool wear prediction results. The relational database is utilized to store metadata about the machine and its attached sensors and data streams. Furthermore, a device management backend interface is used to expose an application programming interface (API) for managing machine metadata and logically associating sensors and data streams. The device management manipulates the relational database mentioned above. Containerizing the software modules via Docker enables flexibility, e. g., for running the pre-processing for multiple machines on one edge device per machine while running model inference on the core infrastructure and running device management and machine frontends in the cloud. Conversely, if there is only one machine connected to the system, all modules can run in a containerized architecture.

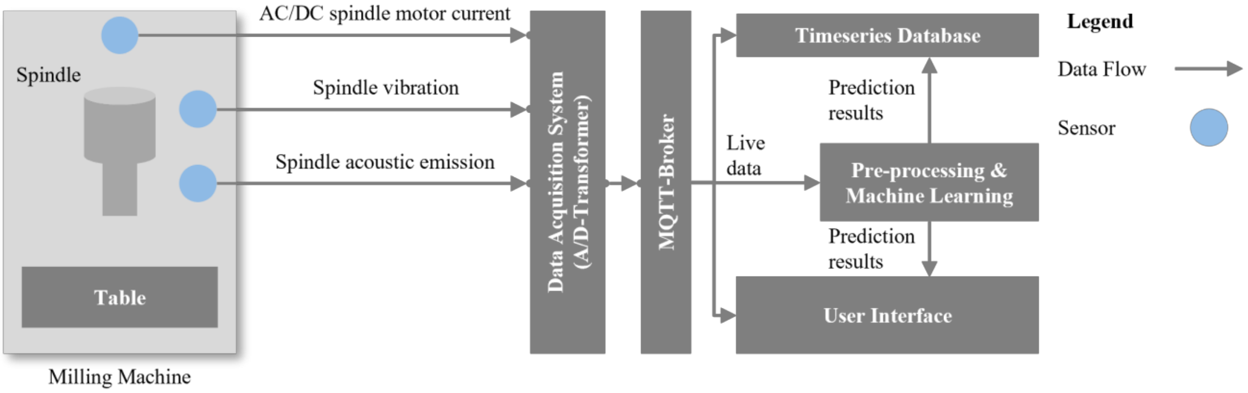


Figure 2: Data flow for the tool wear prediction

The module "Machine Learning Model Tool Wear Prediction" consists of ML data pre-processing and ML models' training and inference. Since the time-domain sensor signal (acoustic emission, vibration, electric motor current) can be well analyzed in the frequency domain and the frequency characteristic changes along time, time-frequency features can be extracted from the raw sensor signal during pre-processing. The extracted features are used as the input to the ML model for prediction purposes. Given the input sensor data, the ML model outputs the tool flank wear which is displayed on the user interface. The application of ML models to a CNC milling machine follows a parent-child methodology. The parent models are pre-trained models based on existing datasets or physical laws. The child models are obtained by further adjusting the parameters of the parent models using the sensor data collected from real-environment CNC machines. Therefore, the child models can be machine and environment-specific, which provide performance improvement over parent models.

The adoption of the parent-child methodology requires two major steps. The first step is offline pre-training, and the second step is online fine-tuning. During the offline pre-training, state-of-the-art ML models are trained using existing data, such as public datasets. Suitable features and suitable models are selected based on their prediction accuracy. These pre-trained models are parent models. Recommended features are time-frequency features, such as those based on wavelet transform and short-time Fourier transform. Recommended models include neural networks, such as multilayer perceptron, 1D convolutional neural network, and recurrent neural network. The parameters of a parent ML model are initialized randomly. Then, given the sensor data from existing datasets, the model outputs a predicted tool flank wear value. A loss function is used to measure the difference between the predicted value and the ground-truth value, to ensure the model parameters are updated towards the direction of minimizing such difference. After installing retrofitting sensors to CNC milling machines, real-time sensor data can be collected, and the online fine-tuning starts. During the online fine-tuning, the parameters of the parent models are adjusted using the sensor data, yielding the child models. The child models require fewer training data as compared to the parent models. Nevertheless, when the system continuously collects sensor data, the child models can be continuously updated to improve performance.

The prediction results are displayed on a dashboard, which enables remote monitoring of multiple machines. To handle large volumes of data and reduce network traffic, the mean data values are calculated and displayed. Figure 3 shows the design of the frontend for one machine (RN-C541). It comprises a dashboard to monitor the cutting tool of the machine connected to the edge device as well as the machine's condition. Both real-time sensor data and real-time predicted tool flank wear are displayed.

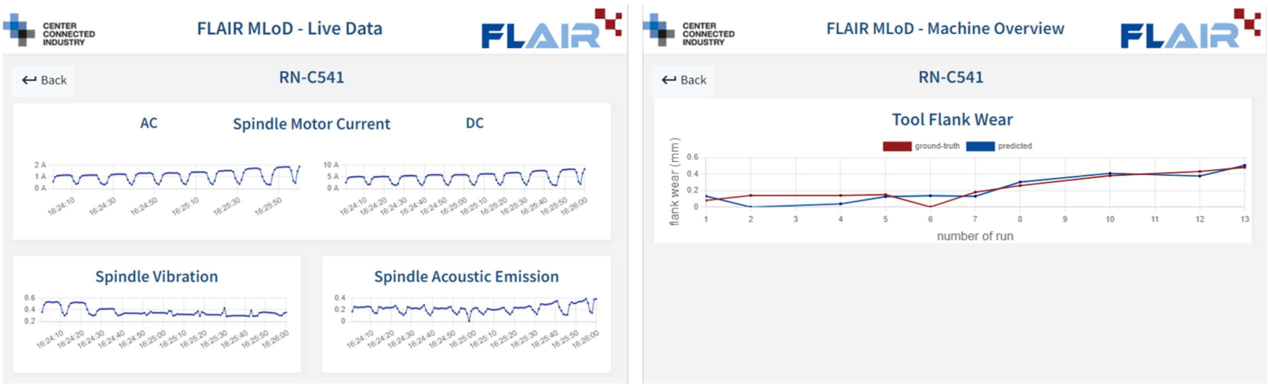


Figure 3: Design of the user interface

3.2 System hardware

Considering the effectiveness of different sensors for tool wear prediction, three types of sensors are selected: the AC current sensor, the vibration sensor, and the acoustic emission sensor. To collect the data accordingly, the current sensor should cover a range of 0 – 100 A. A frequency range of 0 – 20 kHz for the vibration sensor and 50 – 200 kHz for the acoustic emission sensor is required. The vibration sensor and acoustic emission sensor should be waterproof (Ingress Protection IP 68), as they are installed near the spindle and the workpiece, where coolant exists. Based on these requirements, for vibration sensor, the PCB Piezotronics 622B01 accelerometer has been selected, with a sensitivity of 100mV/g and a frequency range of 0.2Hz-15kHz. The vibration sensor can be connected to the edge device through a USB signal conditioner. For acoustic emission sensor, Fujicera AE204SW has been selected, with a sensitivity of 66dB and a frequency range up to 200kHz. A DAQ board is needed to do signal amplification and A/D conversion. The current sensor is a 100A AC current probe. It shares the same DAQ board with the acoustic emission sensor. For edge device, Compulab Tensor-I20 Multi-IoT is selected, with Intel Xeon E-2276ML CPU, 16GB memory, and 1TB storage. It runs Linux operating system, supports RS232/RS485, CANBUS, GPIO, and USB3.1, which should be sufficient for on-site real-time data collection, processing, storage, and transmission.

3.3 Upgrade Kit application methodology

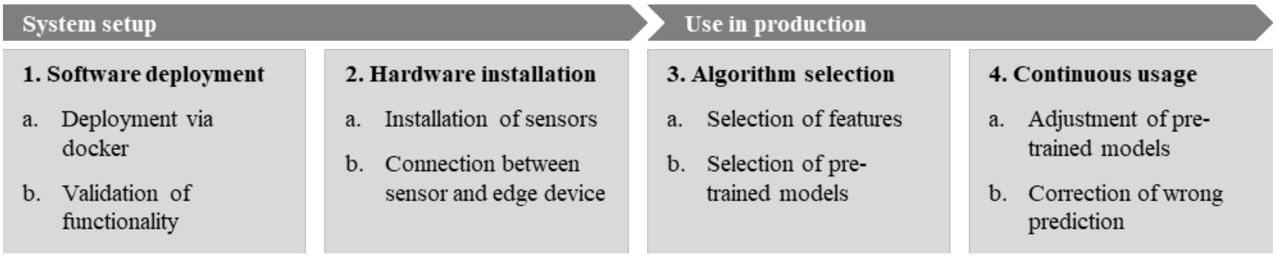


Figure 4: Upgrade Kit application methodology

The application methodology of the Upgrade Kit to real-environment CNC milling machines includes 4 steps. The first step is the deployment of the software on edge devices. The second step is the installation of retrofitting sensors on CNC machines and the communication between sensors and edge device. The third step is the selection of suitable feature extraction algorithms and suitable pre-trained ML models (i.e., parent models). The last step is the validation of the whole system and the continuous collection of data. An overview is given in Figure 4. In the first step, the designed software is deployed as Docker containers to the edge device for a specific CNC machine. Usually, one CNC machine is covered by one edge device, for on-site data collection, storage, processing, and transmission. Nonetheless, adjacent CNC machines may also share the same edge device depending on the specific situation in the shopfloor. In the second step, selected retrofitting sensors, such as current sensor, vibration sensor, and acoustic emission sensor, are installed on the CNC machine. The current probe is clamped on the spindle motor current cable; the acoustic emission sensor is installed on the spindle using a magnetic holder; the vibration sensor requires a screw mounting pad and a clamping fixture. Sensor data are transmitted to the edge device via cables. Data is stored locally and transmitted to the frontend for remote monitoring. The sensor data can be stored on a remote server for long-term recording purposes. In the third step, a default set of features and pre-trained parent ML models are selected. It is optional that the operator selects suitable features and models manually. After determining the features and parent models, real-time data collection can be started. The pre-trained parent models can be directly used to predict tool flank wear values. If newly collected sensor data are paired up with measured tool flank wear values, they can be used to adjust the parameters of the parent model, yielding a child model. In the last step, the whole system is tested for robustness and functionality in a long-term run, such as the integrity of data, the transmission efficiency of data, the proper functioning of the backend (e.g., database) and the frontend (e.g., dashboard). A feedback mechanism is adopted, through which the operator can manually correct the prediction by changing some parameters, such as the offset value.

4. Discussion of the solution

The solution is evaluated in regards to the requirements in section 2.4. The independence of CNC machine vendors (req. a) is achieved by the indirect sensor measurement of the tool wear. Thereby, no interface or API to machine-specific hardware or software is necessary. All hardware and software components are part of the system; thus, a vendor-agnostic approach is realized. The adaptability to company-specific CNC machines (req. b) is achieved by the flexibility of retrofitting sensors and the adaptability of ML models. On the one hand, suitable sensors can be chosen according to specific requirements or necessities of different CNC machines. On the other hand, the parent-child methodology enables further adjusting the parameters of pre-trained parent models to produce the machine-specific child models, using real-time collected machine-specific sensor data. The child models shall provide performance improvements.

The easy deployment (req. c) is realized by the containerized deployment strategy, which avoids cumbersome environment settings on different operating systems and enables easy transferring to multiple machines. The high scalability is achieved by using lightweight IoT protocols, such as MQTT. IoT setups at scale are mainly limited by the available data transmission rate of the utilized network as well as the computational power and storage capacity of the IoT devices. On using MQTT, a publish-subscribe architecture is established, which requires fewer resources for data transmission and makes better use of the network bandwidth.

The system is designed for industrial usage (req. d), by the adoption of industrial-grade sensors and edge device, which support 24-hour continuous operation. It is also designed to be operator-friendly (req. e), by the adoption of remote monitoring. A dashboard continuously displays real-time sensor data and the predicted tool wear. Data are stored locally in the edge device, which can be retrieved for further analysis.

5. Conclusion and future work

The research work outlines a tool wear prediction Upgrade Kit for CNC machines. The system architecture is independent of machine type and vendor. The industrial applicability of the system is reached by industrial-grade edge devices that can be deployed as an addon to legacy machines on the shop floor. The combination of multiple sensors, e.g., electric current, vibration, and acoustic emission for monitoring the machining process, increases prediction accuracy. The scalability of the approach is supported by lightweight IoT protocols and a containerized IT framework, including a backend for data storage and a frontend for data visualization. The adaptability of the approach is achieved by the flexibility in sensor selection and the adjustability of pre-trained ML models. An application methodology is outlined, which enables the application of the Upgrade Kit in diverse industry environments. In the next step, the system will be validated in an industrial environment. Preliminary experiments using an open-source data set [33] show that combining features from multiple sensors outperforms single-sensor features. Therefore, a multi-sensor setup is recommended. In addition, neural networks give better performance than traditional models, such as linear regression and support vector regression. A variety of neural architectures will be included. A feedback mechanism will also be enabled, allowing manual correction on unsatisfactory predictions and manual adjustment on ML model parameters. Data augmentation techniques shall be investigated to reduce the required amount of machine-specific training data.

Acknowledgements

Our research is funded by AIR@InnoHK research cluster of the Innovation and Technology Commission (ITC) of the HKSAR Government. The results presented in this paper within the scope of the research project ‘Machine Learning on Demand for Production Assets’ has come to a successful completion because of the support from ITC. We would like to express our sincere gratitude to them.

The authors want to thank Mansi Patil, Michael Maasem, Peter Gschladt and Marcus Meeßen for their meaningful contribution to the concept and implementation of the solution.

References

[1] Klocke, F., 2018. Verfahren mit rotatorischer Hauptbewegung, in: Klocke, F. (Ed.), Fertigungsverfahren 1. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 525–601.

[2] Wong, S.Y., Chuah, J.H., Yap, H.J., 2020. Technical data-driven tool condition monitoring challenges for CNC milling: a review. The International Journal of Advanced Manufacturing Technology 107 (11-12), 4837–4857.

[3] Goyal, D., Pabla, B.S., 2015. Condition based maintenance of machine tools—A review. CIRP Journal of Manufacturing Science and Technology 10, 24–35.

[4] Ghosh, S., Kumar Naskar, S., Kumar Mandal, N., 2018. Estimation of residual life of a cutting tool used in a machining process. MATEC Web Conf. 192 (11), 1017.

[5] Peng, Y., Dong, M., Zuo, M.J., 2010. Current status of machine prognostics in condition-based maintenance: a review. The International Journal of Advanced Manufacturing Technology 50 (1-4), 297–313.

[6] Brecher, C., Weck, M., 2021. Prozessüberwachung, in: Brecher, C., Weck, M. (Eds.), Werkzeugmaschinen Fertigungssysteme 3, 9. ed. Springer, Berlin, pp. 193–280.

[7] Benker, M., Furtner, L., Semm, T., Zaeh, M.F., 2021. Utilizing uncertainty information in remaining useful life estimation via Bayesian neural networks and Hamiltonian Monte Carlo. Journal of Manufacturing Systems 61, 799–807.

[8] Welte, R., Estler, M., Lucke, D., 2020. A Method for Implementation of Machine Learning Solutions for Predictive Maintenance in Small and Medium Sized Enterprises. Procedia CIRP 93, 909–914.

[9] Hesser, D.F., Markert, B., 2019. Tool wear monitoring of a retrofitted CNC milling machine using artificial neural networks. Manufacturing Letters 19, 1–4.

[10] Lins, R.G., Guerreiro, B., Schmitt, R., Sun, J., Corazzim, M., Silva, F.R., 2017. A novel methodology for retrofitting CNC machines based on the context of industry 4.0: ISSE 2017 Symposium Proceedings. IEEE, 6 pp.

[11] Niemeyer, C.L., Gehrke, I., Müller, K., Küsters, D., Gries, T., 2020. Getting Small Medium Enterprises started on Industry 4.0 using retrofitting solutions. Procedia Manufacturing 45, 208–214.

[12] Sayyad, S., Kumar, S., Bongale, A., Kamat, P., Patil, S., Kotecha, K., 2021. Data-Driven RUL Estimation for Milling Process: Sensors, Algorithms, Datasets, and Future Directions. IEEE Access 9, 110255–110286.

[13] Kuntoğlu, M., Aslan, A., Pimenov, D.Y., Usca, Ü.A., Salur, E., Gupta, M.K., Mikolajczyk, T., Giasin, K., Kapłonek, W., Sharma, S., 2020. A Review of Indirect Tool Condition Monitoring Systems and Decision-Making Methods in Turning: Critical Analysis and Trends. Sensors 21 (1).

[14] Zhou, Y., Xue, W., 2018. A Multisensor Fusion Method for Tool Condition Monitoring in Milling. Sensors 18 (11).

[15] Scheffer, C., Kratz, H., Heyns, P.S., Klocke, F., 2003. Development of a tool wear-monitoring system for hard turning. International Journal of Machine Tools and Manufacture 43 (10), 973–985.

[16] El Hakim, M.A., Shalaby, M.A., Veldhuis, S.C., Dosbaeva, G.K., 2015. Effect of secondary hardening on cutting forces, cutting temperature, and tool wear in hard turning of high alloy tool steels. Measurement 65, 233–238.

[17] Sasahara, H., Satake, K., Takahashi, W., Goto, M., Yamamoto, H., 2017. The effect of oil mist supply on cutting point temperature and tool wear in driven rotary cutting. Precision Engineering 48, 158–163.

[18] Kopač, J., Šali, S., 2001. Tool wear monitoring during the turning process. Journal of Materials Processing Technology 113 (1-3), 312–316.

[19] Kong, D., Chen, Y., Li, N., 2018. Gaussian process regression for tool wear prediction. Mechanical Systems and Signal Processing 104, 556–574.

[20] Luo, B., Wang, H., Liu, H., Li, B., Peng, F., 2019. Early Fault Detection of Machine Tools Based on Deep Learning and Dynamic Identification. IEEE Trans. Ind. Electron. 66 (1), 509–518.

[21] Shi, C., Luo, B., He, S., Li, K., Liu, H., Li, B., 2020. Tool Wear Prediction via Multidimensional Stacked Sparse Autoencoders With Feature Fusion. IEEE Trans. Ind. Inf. 16 (8), 5150–5159.

[22] Traini, E., Bruno, G., D'Antonio, G., Lombardi, F., 2019. Machine Learning Framework for Predictive Maintenance in Milling. IFAC-PapersOnLine 52 (13), 177–182.

[23] Rastegari, A., Archenti, A., Mobin, M. Condition based maintenance of machine tools: Vibration monitoring of spindle units, in: 2017 Annual Reliability and Maintainability Symposium (RAMS). IEEE, pp. 1–6.

[24] Gouarir, A., Martínez-Arellano, G., Terrazas, G., Benardos, P., Ratchev, S., 2018. In-process Tool Wear Prediction System Based on Machine Learning Techniques and Force Analysis. Procedia CIRP 77, 501–504.

[25] Cheng, W.-N., Cheng, C.-C., Lei, Y.-H., Tsai, P.-C., 2020. Feature selection for predicting tool wear of machine tools. The International Journal of Advanced Manufacturing Technology 111 (5-6), 1483–1501.

[26] Traini, E., Bruno, G., Lombardi, F., 2021. Tool condition monitoring framework for predictive maintenance: a case study on milling process. International Journal of Production Research 59 (23), 7179–7193.

[27] Zhou, Y., Sun, W., 2020. Tool Wear Condition Monitoring in Milling Process Based on Current Sensors. IEEE Access 8, 95491–95502.

[28] Xu, W., Miao, H., Zhao, Z., Liu, J., Sun, C., Yan, R., 2021. Multi-Scale Convolutional Gated Recurrent Unit Networks for Tool Wear Prediction in Smart Manufacturing. Chin. J. Mech. Eng. 34 (1).

[29] Cai, W., Zhang, W., Hu, X., Liu, Y., 2020. A hybrid information model based on long short-term memory network for tool condition monitoring. Journal of Intelligent Manufacturing 31 (6), 1497–1510.

[30] Huang, P.-M., Lee, C.-H., 2021. Estimation of Tool Wear and Surface Roughness Development Using Deep Learning and Sensors Fusion. Sensors 21 (16).

[31] Bank, A., Briggs, E., Borgendale, K., Gupta, R., 2019. MQTT Version 5.0. https://mqtt.org/mqtt-specification/.

[32] Jennings, C., Shelby, Z., Arkko, J., Keranen, A., Bormann, C., 2018. Sensor Measurement Lists (SenML)-RFC 8428. RFC Editor.

[33] Agogino, A., Goebel, K., 2007. Milling Data Set. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/.

Biography

Yuechi Jiang received his Ph.D. Degree in Electronic and Information Engineering from Hong Kong Polytechnic University in 2021. He is currently a Postdoctoral Researcher at FLAIR, working on data science.

Benny Drescher serves as Chief Technical Officer at FLAIR and Director Technology at INC, Hong Kong, SAR China. He holds a B.Sc./M.Sc. in Electrical Engineering and a Ph.D. (Dr.-Ing.) in Mechanical Engineering from Technical University of Munich, Germany.

Max Wittstamm received his M.Sc. Degree in Production Engineering at RWTH Aachen University. Currently, he is a Product Owner at FLAIR, working on the transfer of AI to production.

Cuihong Hu received her M.Sc. Degree in Biomedical Engineering from Xi'an Jiaotong University. She is currently a Senior Software Developer at FLAIR, working on architecture design.

Florian Clemens received his M.Sc. Degree in Industrial Engineering at University Paderborn. He is a Researcher at FIR since 2019 in the department Information Management, working on the transfer of AI to production.

Weimin Wang received the Ph.D. Degree in Mechanical Manufacture and Automation from the South China University of Technology. He works on AI applications in Electrical and Mechanical Automation.

Prof. Dr.-Ing. Volker Stich has been CEO of the FIR since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and led the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

Designing A Blockchain-Based Digital Twin For Cyber-Physical Production Systems

Larissa Krämer[1], Nick Große[2],
Rico Ahlbäumer[1], Patrick Stuckmann-Blumenstein[2],
Michael Henke[2], Michael ten Hompel[1]

[1]*Chair of Material Handling and Warehousing, TU Dortmund University, Germany*
[2]*Chair of Enterprise Logistics, TU Dortmund University, Germany*

Abstract

Trust in all processes on the shopfloor is crucial for the success of a production process, especially in cross-company scenarios such as shared manufacturing, in which independent parties interact with each other. A cyber-physical production system (CPPS) contributes to the vision of a decentralized, self-configuring and flexible production. Digital twins (DTs) can visualize the material, information and financial flows in real-time and improve the process transparency of such production systems. The efficiency of digital twins depends on the integrity of the provided data, especially if data is shared across company borders. Due to its characteristics such as immutability and transparency, blockchain technology (BCT) provides a basis for establishing the desired trust in the systems on the shopfloor. This paper proposes the design of a BCT-based DT in CPPS. The design is demonstrated by a prototype including smart contracts attached to a CPPS simulation model visualizing the information and material flow. Tasks are decentrally allocated, deployed and safely documented via blockchain. The demonstrator is revealing supplementary benefits in terms of transparency provided by the BCT. This paper further examines whether BCT can enrich existing solutions and provide a reliable information basis for profound data and process analysis.

Keywords

Cyber-Physical Production System; Blockchain; Negotiation; Digital Twin; Manufacturing; Transaction

1. Introduction

Cyber-physical systems (CPS) are the result of merging virtual and physical entities, whose interplay is enabled by embedded hard- and software attached to its respective physical counterpart [1]. Transferred to the shopfloor, cyber-physical production systems (CPPS) are a promising approach to fulfill the vision of a smart factory [1], which is realized through the interplay of CPS and so-called digital twins (DTs) [2]. However, the visibility enabled by DTs is impeded by restricted data accessibility or lack of data collections [3].

Due to its potential of '*provid*[ing] *validated, immutable transactions, i.e. database updates* […]' [4, p.1546], blockchain technology (BCT) is herein considered. The interplay of its inherent consensus mechanism, cryptographic protocol and distributed storage lays the ground for an immutable and transparent data storage to achieve data and process integrity [4–7]. Blockchain itself can be adopted for '*redesign*[ing] *informational and financial flows, both of which supplement physical flows in a supply chain*' [8, p.10].

DOI: https://doi.org/10.15488/12182

A BCT-based DT promises a reliable real-time visualization of these flows for CPPS. Against that backdrop, the underlying research question is how to design a BCT-based DT for CPPS. For doing this, a methodology for conducting and carrying out Design Science Research (DSR) proposed by [9] is applied here. Our contribution is structured as follows:

- Introducing the addressed problem and objectives (**chapter 1**)
- Deriving the state of the art of DT, BCT, CPPS and its intersections (**chapter 2**)
- Designing a BCT-based DT for CPPS (**chapter 3**)
- Demonstrating and evaluating the prototype (**chapter 4**)
- Communicating design implications (**chapter 5**)

2. State of The Art: Conjunction of CPPS, Digital Twin and Blockchain Technology

The starting point for the design process is the derivation of design knowledge about the three investigation units CPPS, DT and BCT. In each subsection, the investigation unit is introduced, followed by a summary of previous research dealing with the intersection with the other two units. The findings form the design knowledge for the subsequent designing process (**chapter 3)**.

Cyber-Physical Production System (CPPS)

Increasing product variety and new technological developments have inpired the idea of CPPS that allow for automated, flexible and self-configuring production [10]. However, autonomous, decentrally organized CPPS have to deal with high levels of complexity and much uncertainty regarding cross-company interactions as well as issues with data security and robustness against failures [11].

In a CPPS, heterogeneous entities such as mobile robots, smart bins and machines interact with each other. These entities do not necessarily belong to the manufacturer, but can be provided by several independent parties such as vendors or lessors [12]. Transparency with regard to process data (e.g. usage data) is mandatory for a smooth collaboration and reduced coordination effort in terms of payment. Furthermore, as the entities utilize different software, universal interfaces are necessary to ensure a comprehensive and transparent overview across all processes [13]. However, conversing between different interfaces is prone to errors, which could cause loss of information. Besides, unauthorized devices with reading and writing rights could manipulate data such as order data. Without precise, secure and tamper-proof documentation these failures and manipulations affect the transparency of the CPPS for the involved parties [14].

Additionally, third parties such as customers and suppliers call for transparency about the progress of an order or the status of the warehouse for procurement. Trust in the manufacturing process determined by correct and sufficient quality control and punctual delivery is important for a successful customer loyalty [15]. Thus, a lack of transparency can result in inefficient procurement processes and dissatisfied customers [16]. It leads to a higher amount of overhead as costs of robots or smart bins cannot be assigned to the causative principle. Non-transparent process documentations and overheads affect the reduction of process costs and the precise determination of prices and profits of single products [17].

Digital Twin (DT)

To cope with some of these challenges, a DT can be used in a CPPS. The term DT was introduced by [18] in 2003 and describes the precise representation of a physical object or system in the digital (cyber) world [18]. Physical and digital systems can affect each other [18,19], which leads to a merge of the physical system and its digital representation [20]. Further distinctions of DT have been made by [21], in which the DT consists of a bilateral information flow between a physical object and its digital representation [21]. Extensive research of applying DTs into the production field on shop floor level can be traced back to the elaborations of [22,23]. The authors propose a five-dimensional DT structure, which consists of the physical entity, its virtual pendant, services between both entities, data of the DT and the connection of different parts of a DT [23]. Following [24], two starting scenarios for DT development can be distinguished. In the first scenario, neither a physical object nor a DT of such an object exists, whereas the second scenario concentrates on extending an existing physical object that does not have a DT yet. Each scenario passes the design, development, operational and dismissal phase [24].

DTs can be applied for data analysis and simulation to reduce costs, predict failures and prepare for unexpected events. Simulating specific states of the system or adding new components is also possible with the DT without changing the physical system. This saves costs and time, for example when developing a prototype or testing different production scenarios. The DT is also used for 2D or 3D visualisation of production processes [25]. Despite the potential of using a DT in CPPS, several aspects remain critical in a combined system. The DT might not display the physical part of the CPPS correctly due to a lack of transparency and documentation. Besides, as explained above security such as device authorization is a critical issue in a CPPS. Manipulated data of the physical system influences the DT due to the interplay between the physical and digital system such as the use of sensors [25]. Therefore, the DT is highly dependent on reliable data from the physical world, whereas the physical system is dependent on the correct input from the digital world.

Blockchain-Technology (BCT)

One approach to the solution is BCT. According to DIN SPEC 16597, the term blockchain can be described as a '*distributed database that is practically immutable by being maintained by a decentralized P2P* [peer-to-peer] *network using a consensus mechanism, cryptography and back-referencing blocks to order and validate transactions*' [26, p.8]. BCT can be subordinated to the Distributed Ledger Technology (DLT) as a DLT concept [27]. A condensed overview of the most mentionable challenges and benefits related to blockchain is proposed by [28]. Following [29], BCT is characterized by its permanence, immutability, disintermediation and transparency and receives closer consideration due to these trust-inherent characteristics enabled by the interplay of consensus mechanism, decentralization and cryptography [4,7]. An emphasis in terms of establishing trust in intercompany networks enabled by BCT is proposed by [30].

A proposal for a BCT-based engineering framework is described in [5], which consists of technological components on an infrastructure layer, enriched by an environment layer, an application layer, an agent layer, a behavior layer and the trust frontier, which separates the latter ones and addresses the trust issues between the physical system and its virtual model. The interplay of immutability and transparency forms the basis for process and data integrity on the application layer, in which the data integrity consists of the degree of completeness and immutability of data and the process-integrity encompasses the rule-compliant execution of processes [7]. In trust-relating literature, the term integrity describes the '*perception that the trustee adheres to a set of principles that the trustor finds acceptable*' [31, p.719]. In conjunction with the ability and benevolence, it represents one of the '*factors of trustworthiness*' [31, p.717]. The authors of [32] assume that an increase in integrity can be enabled by smart contracts as long as they limit the scope and expectation of opportunistic behavior. Smart contracts can be seen as '*autonomous interacting pieces of code* […]' [4, p.1543], whose execution is based on predefined rules and which allow the omission of intermediary instances, that, in turn, saves transaction costs [4,33].

Referring to [29], '*the process of creating a 'digital twin' of a physical good on a blockchain is called tokenization* [in which] [...] *Users can exchange the ownership of these digital representations, or tokens* […]' [29, p. 3]. DLT allows the virtual representation of the properties or behavior of the reference object, which can be either a person or an object [34]. An example, how a DT can be connected to the Ethereum blockchain is shown in [35], where a three-layered concept is proposed, in which non-fungible (ERC721) Ethereum tokens are connected to DTs and enriched by smart contracts, which '*allow automatic material flow decisions in the manufacturing system*' [35, p. 253].

Towards a DLT-based DT for CPPS

In a previous elaboration on CPPS and DLT, a literature research led to the assumption that DLT enriches previous (de-)centralized infrastructures in terms of security and processing [36]. Furthermore, integrating a BCT-based DT in a CPPS provides a solution for the challenges in such production systems [37,38]. It can be used to interact with the system and carry out payment or replenishment orders. Besides, it does not only present the current state of information in the system, but also offers a complete and tamper-proof history of all states the system has been in so far. Thus, it serves as the basis for sound decision-making and profound process analysis [39] and provides reliable data for AI approaches such as the prediction of failures or peak loads [40]. Furthermore, via BCT and a DT, data exchange within the CPPS and between external stakeholders such as customers and suppliers can be standardized and simplified [41].

The potential of combining CPPS with BCT has been recognized in literature: Smart contracts can provide a distributed and immutable record of transactions [42] by checking for fulfilled requirements and automatically reacting to that. Based on this, pay-per-use models can be realized [43,44]. Several publications either propose theoretical concepts for a BCT-controlled CPPS [38,45,46] or present small-scale implementations of single components of a CPPS [37,47]. These also include, for example, blockchain-based auctions usable for task allocation [48,49]. It is evident from the researched literature that BCT provides transparency in CPPS due to its inherent characteristics. However, the triad of CPPS, BCT and DT has only been researched in few publications on a conceptual level [40,50]. As of now, it remains unclear how to design and implement the BCT-based DT to realize such potential.

3. Designing the BCT-based DT of a CPPS

For designing and implementing a DT, the prerequisites under which the DT is developed and its lifecycle have to be considered. Referring to [24], we focus on the second scenario as our DT extends an existing CPPS in our research hall. The focus is set on the design phase with regard to the DT of an excerpt of a physical production system and its connection to a blockchain. Our design of the DT is based on the five dimensions of a DT proposed by [23]. We combine these dimensions with the blockchain engineering framework proposed by [5] with intention to overcome the challenges of process and data integrity. The merged model, transferred to CPPS, is shown in **Figure 1.**

The **behavioral layer** consists of a tangible CPPS with its entities such as machines and mobile robots and represents the DT dimension of the Physical Entity. This tangible CPPS is represented on the **agent layer** by one or several DTs (**virtual entity**) enabled by simulation tools [51]. The trust issues between the agent layer and the behavioral layer can be described by the lack of process and data integrity as trust-inherent characteristics [7]. Blockchain can be attached to the DT of the CPPS and supports the visibility of the state conditions enabled by smart contracts on the **application layer** [4,5]. The interplay and connection of the physical and virtual entity as well as the blockchain represents the dimension of the **connected digital twin**, in which information are exchanged in a bidirectional way between the layers.

The application layer itself ensures the interoperability between the blockchain framework and the agents, which can be either a human frontend interaction or a machine-API-interaction of a cyber-physical entity acting as an agent. Transferred to the scenario of a CPPS, smart contracts can be foreseen for executing the transactions inside the negotiation scenario on the shop floor. This drives the automatizing process forward and allows agents to act autonomously. These interactions are part of the DT dimension called **services,** which also includes the interoperability between the physical and virtual entity by synchronizing them in real-time. The physical entity can report its initial states of the system to the virtual entity and the virtual entity can in turn influence the **physical entity**, for example by returning a determined resource allocation. Other services, for instance, can be provided for monitoring of energy consumption in the physical entity or testing of certain functions in the virtual model.

On the **environmental layer**, the policies of the whole framework can be defined, which can be for example rules for task allocation, the definition of roles with certain access rights and restrictions, transaction relevant tokens and the determination of data captured and stored within the transaction log. This is part of the **digital twin data,** which comprises all data from the physical and virtual entities and services and, thus, includes information from every layer. For the CPPS, data from the information, material and financial flows can be part of the DT data. Especially the financial transactions extend the data of the physical system and its virtual counterpart. These include, for example, tenders, bids, order placements and payments.

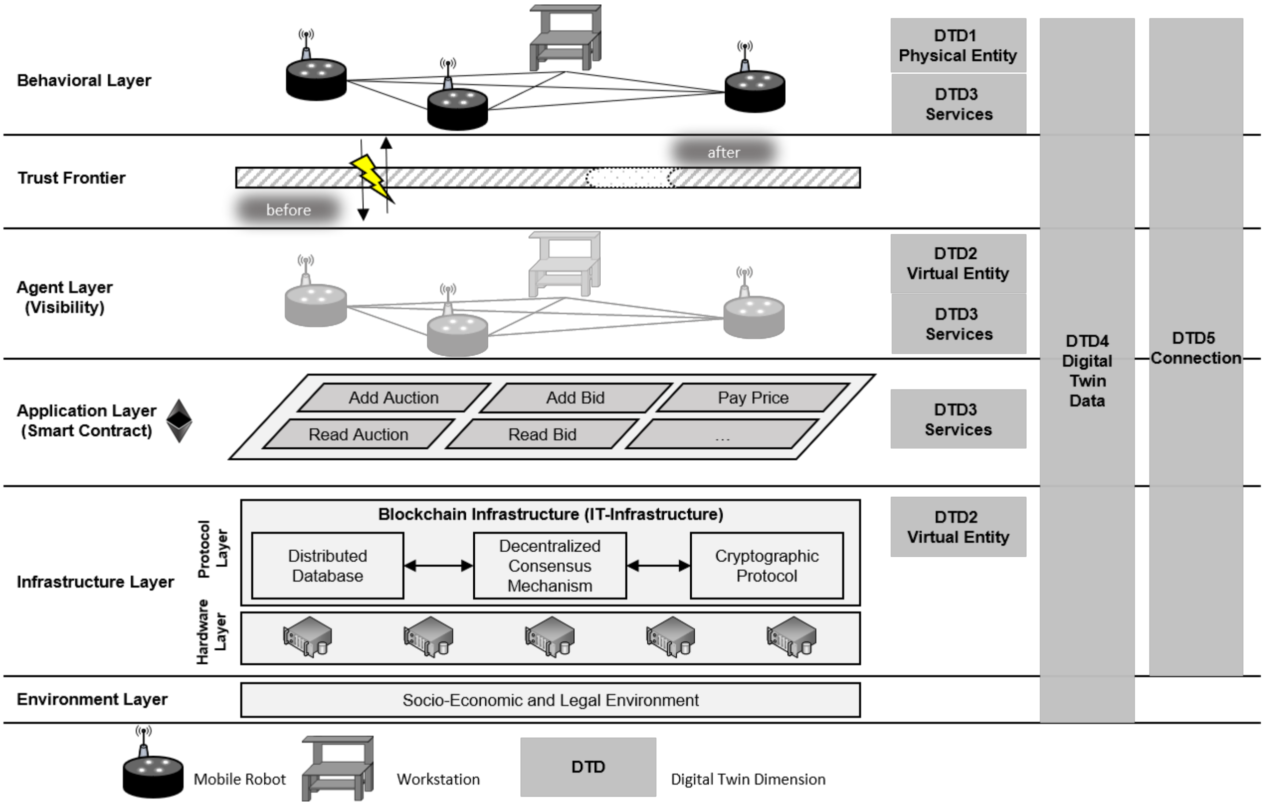


Figure 1: Framework for designing a BCT-based DT for CPPS (based on [5,23])

4. Demonstration and Evaluation

The BCT-based DT of a CPPS to be demonstrated in this contribution is located in the manufacturing context as described by [52,45,53]. In this tangible CPPS, individually configurable drones are manufactured in a decentrally organized matrix production system [54]. The system consists of several entities such as workstations, mobile robots, smart bins and workers that solve task allocation by negotiation. The prototype comprises an excerpt of this production system. The prototype illustrated in **Figure 2** focuses on the task allocation between a workstation and several mobile robots, which provide material for the workstation.

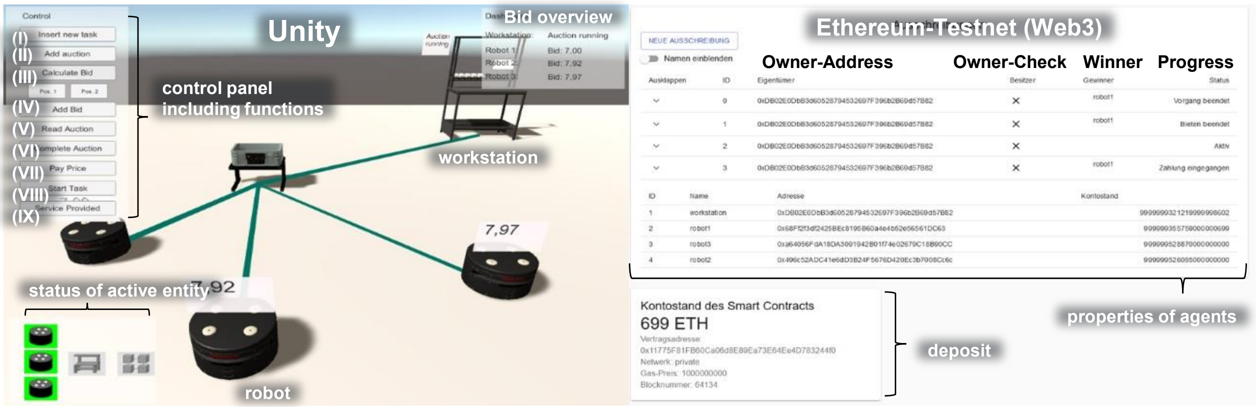
In the following, the prototype considers the agent and application layer with regard to the virtual entity, services, DT data and the connection between the layers. The agent layer of the prototype is replicated with the game engine Unity[1], which also serves as the virtual entity. For visualization purposes, the DT is created and animated in 3D. Inside the Unity scene, an additional 2D dashboard depicts system information such as the current state of a mobile robot in real-time. The virtual entity is supplemented by a blockchain testnet based on the Ethereum framework[2]. The transaction data is stored inside a blockchain and shown in a

---

[1] https://unity.com/

[2] https://ethereum.org/en/

frontend layer, which is provided by a Web3-React interface[3]. Unity and Ethereum are merged via Nethereum[4], an open-source API for integrating the Ethereum blockchain into .NET applications.



Ausklappen	ID	Eigentümer	Besitzer	Gewinner	Status
⌄	0	0xDB02E0Db83d60528794532697F396b2B69d57B82	✕	robot1	Vorgang beendet
⌄	1	0xDB02E0Db83d60528794532697F396b2B69d57B82	✕	robot1	Bieten beendet
⌄	2	0xDB02E0Db83d60528794532697F396b2B69d57B82	✕		Aktiv
⌄	3	0xDB02E0Db83d60528794532697F396b2B69d57B82	✕	robot1	Zahlung eingegangen

ID	Name	Adresse	Kontostand
1	workstation	0xDB02E0DbB3d60528794532697F396b2B69d57B82	9999999321219999998602
2	robot1	0x68Ff2f3df2425BEc8195B60a4e4b52e56561DC63	999999355759000000699
3	robot3	0xa64056FdA18DA3091942B01f74e02679C18B90CC	999999528870000000000
4	robot2	0x496c52ADC41e6dD3B24F5676D420Ec3b7008Cc6c	999999526085000000000

Kontostand des Smart Contracts
699 ETH
Vertragsadresse:
0x11775F81FB60Ca06d8E89Ea73E64Ee4D783244f0
Netwerk: private
Gas-Preis: 1000000000
Blocknummer: 64134

Figure 2: Sketch of the prototype of the merged visualization with unity (left) and blockchain (right)

The prototype is embedded in a scenario comprising the request of material, followed by settlement and task allocation, task fulfillment and payment. The decentralized task allocation follows the principle of rule-based bidding with a reverse first-price-sealed-bid auction (FPSB), in which multiple sellers compete for a single buyer without knowing the bids of the competitors [49]. In this context, robots are offering their transportation services, which are requested by a workstation. For simplicity, the bid of each robot is determined by its distance to the workstation. The workstation chooses the lowest bid, which is the closest robot. Smart contracts are based on the programming language Solidity[5]. As **Table 1** shows they provide call- and send-functionalities for the control of the CPPS.

This prototype shows the relevance of the DT in CPPS with integration of BCT. In our system, the DT based on a virtual entity in Unity allows an extensive visualization of the production system as well as its material and information flows. A real-time environment provides the user with an overview of the whole system at any time to make well-founded decisions. Furthermore, the visual interface allows a human-machine-interaction with the virtual entity. This is a decisive step for successive testing and understanding the cause-effect relationships of DT during the development phase and it is also a mandatory part for the basic DT archetype [51]. Thus, the prototype can be classified as an '*Enriched Digital Twin*', which '*enriches its database by preprocessed data from supplementary systems* […]' [51, p.12].

Additionally, the Unity model allows an analysis in time-lapse without the need of validating transactions in a blockchain. Physical behavior can be replicated and negotiation-relevant data such as the bids of the robots can be calculated faster. The virtual entity enabled by the Ethereum framework provides the requested transparency and process as well as data integrity. As it would not be efficient to store all data on a blockchain due to the necessary computing power and storage, only data with high need of integrity has to be stored on the blockchain [46]. Thus, the blockchain does not replicate the holistic DT in Unity. It rather enriches it by providing several functions to ensure compliance and transparency and promotes the virtual entity to the next level of an enriched DT archetype [51].

The Web3-frontend and the blockchain allow for selective control of access rights in the CPPS. Only verified entities can enter the system and add auctions or bid on auctions, which means that they have to belong to the requested group of entities or that they have to have a certain amount of Ether to be able to pay for a tender. These features are considered in the provided smart contracts, for example the `addBid` function.

---

[3] https://betterprogramming.pub/blockchain-introduction-using-real-world-dapp-react-solidity-web3-js-546471419955

[4] http://docs.nethereum.com/en/latest/

[5] https://docs.soliditylang.org/

Requirement statements embedded in the functions prevent entities from bidding on their own auctions and check if the entities fulfil the requirements such as necessary rights to call these functions. Hence, rule compliant negotiation mechanisms within the CPPS can be ensured and process integrity can be guaranteed. As all transactions are recorded safely and completely, they can be used to automatically assign costs according to the causative principle and to ensure that the respective payments are completed. If an auction is completed and the winner is selected, the owner of the auction will deposit the amount of the bid in Ether in the smart contract. The winner is informed and fulfils its task. Once the task has been fulfilled, the smart contract automatically releases the stored amount of Ether as payment for the winner. This measure creates trust for both entities, as the winner has certainty about the fulfilled payment, whereas the owner has confidence that it will get its Ether back in case of failure. The blockchain keeps a complete and tamper-proof history of transactions including bidding information, owners of auctions, selected winners and deposits, which persists over the limited runtime of the Unity model.

Table 1: Overview of functionalities

No.	Function	Input Parameters for Smart Contracts	Supported by virtual entity
(I)	`insertNewTask`	-	Unity
(II)	`addAuction`	`(privateKey)`	Ethereum
(III)	`calculateBid`	-	Unity
(IV)	`addBid`	`(privateKey, auctionNumber, bid)`	Ethereum
(V)	`readAuction`	`(privateKey, auctionNumber, bid)`	Ethereum
(VI)	`completeAuction`	`(privateKey, auctionNumber)`	Ethereum
(VII)	`payPrice`	`(privateKey, auctionNumber)`	Ethereum
(VIII)	`startTask`	-	Unity
(XI)	`serviceProvided`	`(privateKey, auctionNumber)`	Ethereum

5. Conclusion and outlook

This contribution focuses on the triad of CPPS, BCT and DT to enhance transparency, and process and data integrity in CPPS. A concept of a BCT-based DT for CPPS is designed and demonstrated. Based on reusable design knowledge about each component, a consolidated framework for designing BCT-based DT for CPPS is proposed. The demonstration of a prototype of a BCT-based DT for CPPS in the manufacturing scenario confirms that the transparency and integrity within the system can be increased. However, the elaboration underlies certain limitations. More profound research of additional challenges concerned with BCT such as security, privacy or scalability is required [28], especially with CPPS and DT. The methodology followed allows further revisions during the build and evaluate cycle [9]. In future, we will extend the presented prototype to the whole system in terms of the scope and variety of agents and corresponding transactions. This includes a more resilient quantitative analysis regarding validation time and transaction costs as well as scalability. Other open aspects are the integration of the physical entity in the development phase and the elevation of the DT to a higher level of archetype by integrating autonomous BCT-based control of the extensive CPPS [51]. The broadly accepted value propositions of BCT and DLT as for example trustworthiness, credibility, immutability, data sovereignty or decentralization pave the way for a next evolution level adding a Financial DT component to the concepts of DTs of CPPS on the shopfloor. Financial DTs acting as agents of a company´s Finance Department and interacting with their digital counterparts on the shopfloors might have a huge potential to trigger innovative and highly efficient working capital and asset finance solutions. We encourage researchers and practitioners in the fields of Information Systems and Logistics to adapt the proposed concept to derive design implications in the fields of BCT, DT and CPPS.

Acknowledgement

This research was provided by the ‘Blockchain Europe Initiative' funded by the Ministry of Economic Affairs Innovation, Digitalization and Energy of the State of North Rhine-Westphalia (MWIDE) (Grant No. 005- 2003-0068). The work presented in this paper was funded by the ‘Deutsche Forschungsgemeinschaft (DFG) - 276879186/GRK2193’.

References

[1] acatech - Deutsche Akademie der Technikwissenschaften, 2011. Cyber-Physical Systems: Driving force for innovation in mobility, health, energy and production. Springer, Berlin.

[2] Tao, F., Qi, Q., Wang, L., Nee, A., 2019. Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. Engineering 5 (4), 653–661.

[3] Schuh, G., Anderl, R., Gausemeier, Krüger, A., Hompel, M. ten. Industrie 4.0 Maturity Index: Managing the Digital Transformation of Companies. Update 2020. https://www.acatech.de/publikation/industrie-4-0-maturity-index-update-2020/.

[4] Glaser, F. Pervasive Decentralisation of Digital Infrastructures: A Framework for Blockchain enabled System and Use Case Analysis. Proceedings of 50th Annual Hawaii International Conference on System Sciences (HICCS), 1543–1552.

[5] Hawlitschek, F., Notheisen, B., Teubner, T., 2018. The limits of trust-free systems: A literature review on blockchain technology and trust in the sharing economy. Electronic Commerce Research and Applications 29, 50–63.

[6] Notheisen, B., Hawlitschek, F., Weinhardt, C., 2017. Breaking down the blockchain hype-towards a blockchain market engineering approach. Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10, (pp. 1062-1080).

[7] Wieninger, S., 2020. Vertrauen in Unternehmensnetzwerken durch Blockchain-Technologie. PhD Thesis. Apprimus Wissenschaftsverlag, Aachen.

[8] Treiblmaier, H., 2019. Combining Blockchain Technology and the Physical Internet to Achieve Triple Bottom Line Sustainability: A Comprehensive Research Agenda for Modern Logistics and Supply Chain Management. Logistics 3 (1), 10.

[9] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A design science research methodology for information systems research. Journal of management information systems 24 (3), 45–77.

[10]Monostori, L., 2014. Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. Procedia CIRP 17, 9–13.

[11]Lee, J., Azamfar, M., Singh, J., 2019. A blockchain enabled Cyber-Physical System architecture for Industry 4.0 manufacturing systems. Manufacturing Letters 20, 34–39.

[12]Xia, T., Sun, B., Chen, Z., Pan, E., Wang, H., Xi, L., 2021. Opportunistic maintenance policy integrating leasing profit and capacity balancing for serial-parallel leased systems. Reliability Engineering & System Safety 205, 107233.

[13]Xiao, G., Guo, J., Da Xu, L., Gong, Z., 2014. User Interoperability With Heterogeneous IoT Devices Through Transformation. IEEE Trans. Ind. Inf. 10 (2), 1486–1496.

[14]Suvarna, M., Yap, K.S., Yang, W., Li, J., Ng, Y.T., Wang, X., 2021. Cyber–Physical Production Systems for Data-Driven, Decentralized, and Secure Manufacturing—A Perspective. Engineering 7 (9), 1212–1223.

[15]Čater, T., Čater, B., 2010. Product and relationship quality influence on customer commitment and loyalty in B2B manufacturing relationships. Industrial Marketing Management 39 (8), 1321–1333.

[16]Francisco, K., Swanson, D., 2018. The Supply Chain Has No Clothes: Technology Adoption of Blockchain for Supply Chain Transparency. Logistics 2 (1), 2.

[17]Klotz, L., Horman, M., Riley, D., Bechtel, J., 2009. Process transparency for sustainable building delivery. International Journal of Sustainable Engineering 2 (4), 298–307.

[18]Grieves, M., 2015. Digital Twin: Manufacutring Excellence through Virtual Factory Representation: Whitepaper.

[19]Grieves, M., Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems, in: Kahlen, F.-J., Flumerfelt, S., Alves, A. (Eds.), Transdisciplinary Perspectives on Complex Systems. Springer International Publishing, Cham, pp. 85–113.

[20]Guo, D., Ling, S., Li, H., Ao, D., Zhang, T., Rong, Y., Huang, G.Q., 2020. A framework for personalized production based on digital twin, blockchain and additive manufacturing in the context of Industry 4.0, in: 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE), Hong Kong, Hong Kong. 8/20/2020 - 8/21/2020. IEEE, Piscataway, NJ, pp. 1181–1186.

[21]Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51 (11), 1016–1022.

[22]Tao, F., Zhang, M., 2017. Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing. IEEE Access 5, 20418–20427.

[23]Tao, F., Zhang, M., Nee, A., 2019. Digital twin driven smart manufacturing. Academic Press.

[24]Barricelli, B.R., Casiraghi, E., Fogli, D., 2019. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. IEEE Access 7, 167653–167671.

[25]Park, H., Easwaran, A., Andalam, S., 2019. Challenges in Digital Twin Development for Cyber-Physical Production Systems, in: Chamberlain, R., Taha, W., Törngren, M. (Eds.), Cyber Physical Systems. Model-Based Design. 8th International Workshop, CyPhy 2018, and 14th International Workshop, WESE 2018, Turin, Italy, October 4–5, 2018, Revised Selected Papers, vol. 11615, 1st ed. 2019 ed. Springer, Cham, pp. 28–48.

[26]DIN SPEC 16597, 2018. Terminology for blockchains;: Text in English. Beuth Verlag, Berlin.

[27]Kannengießer, N., Lins, S., Dehling, T., Sunyaev, A., 2019. What does not fit can be made to fit! Trade-offs in distributed ledger technology designs, in: Proceedings of the 52nd Hawaii international conference on system sciences, pp. 7069–7078.

[28]Treiblmaier, H., 2019. Toward More Rigorous Blockchain Research: Recommendations for Writing Blockchain Case Studies. Front. Blockchain 2.

[29]Schmidt, C.G., Wagner, S.M., 2019. Blockchain and supply chain relations: A transaction cost theory perspective. Journal of Purchasing and Supply Management 25 (4), 100552.

[30]Grosse, N., Guerpinar, T., Henke, M., 2021. Blockchain-Enabled Trust in Intercompany Networks Applying the Agency Theory, in: , 2021 3rd Blockchain and Internet, pp. 8–14.

[31]Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. Academy of management review 20 (3), 709–734.

[32]Mehrwald, P., Treffers, T., Titze, M., Welpe, I.M., 2019. Application of Blockchain Technology in the Sharing Economy: A Model of Trust and Intermediation, in: Proceedings of the 52nd Hawaii International Conference on System Sciences, pp. 4585–4594.

[33]Szabo, N., 1996. Smart contracts: building blocks for digital markets. EXTROPY: The Journal of Transhumanist Thought,(16) 18 (2).

[34]Fridgen, G., Guggenberger, N., Hoeren, T., Prinz, W., Urbach, N., Baur, J., Brockmeyer, H., Gräther, W., Rabovskaja, E., Schlatt, V., Schweizer, A., Sedlmeir, J., Wederhake, L., 2019. Chancen und Herausforderungen von DLT (Blockchain) in Mobilität und Logistik, Berlin.

[35]Nielsen, C.P., da Silva, E.R., Yu, F., 2020. Digital Twins and Blockchain – Proof of Concept. Procedia CIRP 93, 251–255.

[36]Große, N., Leisen, D., Gürpinar, T., Forsthövel, R.S., Henke, M., Hompel, M. ten, 2020. Evaluation of (De-) Centralized IT technologies in the fields of Cyber-Physical Production Systems, in: P. Nyhuis, D. Herberger, M. Hübner: Conference on Production Systems and Lgistics (Ed.), Proceedings of the 1st Conference on Production Systems and Logistics. publish-Ing., pp. 377–386.

[37]Afanasev, M.Y., Fedosov, Y.V., Krylova, A.A., Shorokhov, S.A., 2018 - 2018. An application of blockchain and smart contracts for machine-to-machine communications in cyber-physical production systems, in: 2018 IEEE Industrial Cyber-Physical Systems (ICPS). 2018 IEEE Industrial Cyber-Physical Systems (ICPS), St. Petersburg. 15.05.2018 - 18.05.2018. IEEE, pp. 13–19.

[38]Afanasev, M.Y., Krylova, A.A., Shorokhov, S.A., Fedosov, Y.V., Sidorenko, A.S., 2018. A Design of Cyber-physical Production System Prototype Based on an Ethereum Private Network, in: 22nd Conference of Open, pp. 3–11.

[39]Ding, K., Chan, F.T., Zhang, X., Zhou, G., Zhang, F., 2019. Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors. International Journal of Production Research 57 (20), 6315–6334.

[40]Lee, J., Azamfar, M., Singh, J., Siahpour, S., 2020. Integration of digital twin and deep learning in cyber-physical systems: towards smart manufacturing. IET Collaborative Intelligent Manufacturing 2 (1), 34–36.

[41]Gries, S., Meyer, O., Wessling, F., Hesenius, M., Gruhn, V., 2018. Using Blockchain Technology to Ensure Trustful Information Flow Monitoring in CPS, in: 2018 IEEE 15th International Conference on Software Architecture companion. ICSA-C 2018 : proceedings : 30 April-4 May 2018, Seattle, Washington. 2018 IEEE International Conference on Software Architecture Companion (ICSA-C), Seattle, WA. 4/30/2018 - 5/4/2018. IEEE, Piscataway, NJ, pp. 35–38.

[42]Smirnov, A., Teslya, N., 2018. Robot Interaction Through Smart contract for Blockchain-Based Coalition Formation, in: Chiabert, P., Bouras, A., Noël, F., Ríos, J. (Eds.), Product Lifecycle Management to Support Industry 4.0. Springer International Publishing, Cham, pp. 611–620.

[43]Gong, X., Liu, E., Wang, R., 2020. Blockchain-Based IoT Application Using Smart Contracts: Case Study of M2M Autonomous Trading, in: 2020 5th International Conference on Computer and Communication Systems, Shanghai, China. 5/15/2020 - 5/18/2020. IEEE Press, Piscataway, NJ, pp. 781–785.

[44]Ranganthan, V.P., Dantu, R., Paul, A., Mears, P., Morozov, K., 2018. A Decentralized Marketplace Application on the Ethereum Blockchain, in: 4th IEEE International Conference on Collaboration and Internet Computing, Philadelphia, PA. 10/18/2018 - 10/20/2018. IEEE, Piscataway, NJ, pp. 90–97.

[45]Bayhan, H., Schulze Forsthövel, R., Kaiser, P., Hompel, M. ten, 2020. Blockchainbasierte cyberphysische Produktionssysteme. Logistics Journal: Proceedings 2020 (12).

[46]Krämer, L., Ahlbäumer, R., Leveling, J., Detzner, P., Brehler, M., Hompel, M. ten, 2021. Towards a Concept for Blockchain-based Cyber-Physical Production systems. Logistics Journal : Proceedings 2021 (17).

[47]Angrish, A., Craver, B., Hasan, M., Starly, B., 2018. A Case Study for Blockchain in Manufacturing: “FabRec”: A Prototype for Peer-to-Peer Network of Manufacturing Nodes. Procedia Manufacturing 26, 1180–1192.

[48]Braghin, C., Cimato, S., Damiani, E., Baronchelli, M., 2020. Designing Smart-Contract Based Auctions, in: Yang, C.-N., Peng, S.-L., Jain, L.C. (Eds.), Security with Intelligent Computing and Big-data Services, vol. 895. Springer International Publishing, Cham, pp. 54–64.

[49]Omar, I.A., Hasan, H.R., Jayaraman, R., Salah, K., Omar, M., 2021. Implementing decentralized auctions using blockchain smart contracts. Technological Forecasting and Social Change 168, 120786.

[50]Zhang, C., Zhou, G., Li, H., Cao, Y., 2020. Manufacturing Blockchain of Things for the Configuration of a Data- and Knowledge-Driven Digital Twin Manufacturing Cell. IEEE Internet Things J. 7 (12), 11884–11894.

[51]van der Valk, H., Haße, H., Möller, F., Otto, B., 2021. Archetypes of Digital Twins. Bus Inf Syst Eng.

[52]Bayhan, H., Meißner, M., Kaiser, P., Meyer, M., Hompel, M. ten, 2020. Presentation of a novel real-time production supply concept with cyber-physical systems and efficiency validation by process status indicators. Int J Adv Manuf Technol 108 (1-2), 527–537.

[53]Borst, D., Ratke, M., Bayhan, H., Hompel, M. ten, 2019. Feldstudie zu einer FTS-Auftragsvergabe für die dezentral gesteuerte Produktionsversorgung in cyber-physischen Produktionssystemen. Logistics Journal : Proceedings 2019 (12).

[54]Blesing, C., Luensch, D., Stenzel, J., Korth, B., 2017. Concept of a Multi-agent Based Decentralized Production System for the Automotive Industry, in: Advances in Practical Applications of Cyber-Physical Multi-Agent Systems: The PAAMS Collection, Cham. 2017. Springer International Publishing, Cham, pp. 19–30.

Biography

Larissa Krämer, M.Sc. M.Sc. (*1996) is a research assistant at the Chair of Material Handling and Warehousing, TU Dortmund University and an associated member of the research training group adaption intelligence of factories in complex and dynamic environment (GRK 2193). E-Mail: larissa.kraemer@tu-dortmund.de

Nick Große, M.Sc. (*1992) is a research assistant at the Chair of Enterprise Logistics, TU Dortmund University and a member of the research training group adaption intelligence of factories in complex and dynamic environment (GRK 2193). E-Mail: nick.grosse@tu-dortmund.de

Rico Ahlbäumer, M.Sc. (*1992) is a research assistant at the Chair of Material Handling and Warehousing, TU Dortmund University. E-Mail: rico.ahlbaeumer@tu-dortmund.de

Patrick Stuckmann-Blumenstein, B.Sc. (*1999) is a student assistant at the Chair of Enterprise Logistics, TU Dortmund University. E-Mail: patrick.stuckmann@tu-dortmund.de

Prof. Dr. Michael Henke (*1971) is the head of the Chair of Enterprise Logistics, TU Dortmund University and a managing director of the Fraunhofer Institute for Material Flow and Logistics.

Prof. Dr. Dr. h. c. Michael ten Hompel (*1958) is the head of the Chair of Material Handling and Warehousing, TU Dortmund University and a managing director of the Fraunhofer Institute for Material Flow and Logistics.

3rd Conference on Production Systems and Logistics

Framework for the Application of Industry 4.0 in Lithium-Ion Battery Cell Production

Jessica Schmied[1], Alexander Puchta[2], Timon Scharmann[3,4], Hans-Christoph Töpper[5,6],
Achim Kampker[1], Jürgen Fleischer[2], Klaus Dröder[3,4], Rüdiger Daub[6]

[1]Production Engineering of E-Mobility Components (PEM), RWTH Aachen University, Bohr 12, 52072 Aachen, Germany
[2]wbk Institute of Production Science, Karlsruhe Institute of Technology, Kaiserstrasse 12, 76131 Karlsruhe, Germany
[3]Technische Universität Braunschweig, Institute of Machine Tools and Production Technology, Langer Kamp 19b, 38106 Braunschweig, Germany
[4] Technische Universität Braunschweig, Battery LabFactory Braunschweig, Langer Kamp 8, 38106 Braunschweig, Germany
[5]TUMint. Energy Research GmbH, Lichtenbergstraße 4, 85748 Garching, Germany
[6]Institute of Machine Tools and Industrial Management, Technical University of Munich, Boltzmannstraße 15, 85748 Garching, Germany

Abstract

The application of Industry 4.0 in lithium-ion battery cell production enables companies to achieve increased product quality and global competitiveness, since the majority of value creation takes place in this process. Studies have shown, that improving production performance is the most effective way for battery cell manufacturers to become competitive in the increasingly globalized market. To achieve operational excellence, battery manufacturers must adopt the concepts of networked and digitized production. However, holistically introducing digitalization, data systems and Industry 4.0 methods in all sectors of lithium-ion battery cell production currently poses a major challenge as comprehensive approaches are not available. Therefore, a tailored methodology for the evaluation of suitability and introduction of digitalization and Industry 4.0 is presented. The approach addresses all production-related sectors from logistics to plant engineering to quality management via so called application areas. Multiple development stages divide these into the maturity levels in terms of Industry 4.0. To design each application area and stage, Industry 4.0 use cases from battery cell producers, plant manufacturers, and battery-related research projects are clustered and abstracted for general accessibility. It is shown, that abstracted application areas may be assigned either to all production sectors such as communication or to specific fields such as quality methods. Based on the application areas, corresponding toolboxes are established forming the core of a digitalization guide. To increase the level of maturity with regard to Industry 4.0, the presented paper aims at enabling companies to apply appropriate tools from the toolbox to their production. The systematic and efficient development and implementation of digitalization as well as the holistic assessment of a company's maturity are enabled and provide an essential tool towards increased competitiveness.

Keywords

Lithium-ion Battery Cell Production; Competitiveness; Digitalization; Industry 4.0; Production Planning

DOI: https://doi.org/10.15488/12168 publish-Ing.

1. Introduction and state of the art

Developing energy storage systems to meet the growing global demand for storage capabilities sustainably providing electrical energy is one of the key challenges of our time. The development of a variety of storage technologies is driven forward, among which the lithium-ion battery cell plays a key role. Its wide range of advantageous properties has resulted in its application in a wide range of products, among which the electric vehicle is the most prominent representative. Despite the rapid development of the lithium-ion battery, it is still facing technological challenges ranging from the improvement of key characteristics such as energy density and safety, elaborating possibilities for recycling and especially establishing cost-effective mass production [1]. The latter in particular plays an important role concerning market penetration. To overcome this challenge, the holistic digitalization of factories and implementation of Industry 4.0 methods can assist [2]. In various industries, a significant increase in competitiveness through improvements such as reduced downtime, increase in throughput, and lower energy consumption have already been shown. In addition, companies are enabled to collect process and product data along the entire value chain and evaluate it using intelligent algorithms enabling product and process development [3,4].

The European lithium-ion battery production capacity is expected to increase to a 16.8 % share of global production capacity in the next decade [5]. Battery cell producers are relying on the introduction of digitalization and Industry 4.0 methods to sustainably increase their productivity and compete in the highly contested market. In Germany, Varta AG, and Daimler AG are increasing their efficiency and flexibility [6,7], Switzerland-based iQ Power Licensing is aiming at shorter throughput time and lower energy consumption, Northvolt is holistically implementing Industry 4.0 approaches in Sweden [8], Freyr Battery introduced digitals twins of all processes in Norway [9], Tesla, Inc. uses holistic digitalization paired with automation in the USA [10], and Contemporary Amperex Technology Co. Limited is doing the same in China [11]. However, since the implementation is a complex interdisciplinary problem, companies face several challenges ranging from the selection of suitable sensors, their networking to the establishment of automated communication and autonomous process control [12,13]. In addition, the maturity determination regarding Industry 4.0 cannot be performed and a clear vision does not exist. Therefore, the efficient introduction of digitalization and Industry 4.0 requires a systematic, holistic and intuitive approach [2].

Since the topic of digitalization and Industry 4.0 in battery cell production is still new to the field of research, only a few approaches identifying and addressing the arising challenges exist. Data acquisition and evaluation, as well as cyber-physical systems in cell production, were analysed [14,15,16], and a holistic, data-driven approach to battery cell manufacturing was demonstrated [17]. In addition, data mining methods were applied to battery cell manufacturing [18] and digitalization conceptually addressed regarding traceability, digital twins, and end-of-life prediction [19]. It can thus be concluded that there is currently no guide for supporting companies in the holistic introduction of digitalization and Industry 4.0 in battery cell production. Consequently, the potential of the next level of industrialization cannot be accessed efficiently and intuitively to competitively manufacture lithium-ion battery cells.

Thereby, all production-related aspects such as process technology of the individual manufacturing processes, the associated machine and plant technology, and quality management to ensure the required quality must be addressed. Furthermore, organizational aspects of production planning, control and logistics must also be included for a holistic view of Industry 4.0 in battery cell production. Although there are extensive studies and guidelines on battery cell production [20,21] and digitalization [22,23], there is no known combination of the two with a holistic view of all aspects of production technology. Digitalization and Industry 4.0 in battery cell production have been named as a necessary technology breakthrough ("Red Brick Wall") [20]. Therefore, the development of the framework and corresponding toolboxes for the digitalization of battery cell production is the goal of this paper and is intended to make a fundamental and sustainable contribution to competitive battery cell production in Germany. So-called toolkits represent one component of the digitalization guide. New functions in the sense of Industry 4.0 and the Maturity Index according to SCHUH ET AL. are to be derived from the toolkits for Industry 4.0 and successfully implemented [24,25]. How this can be achieved is explained in this contribution.

2. Methodology for the development of toolboxes for digitization

This chapter refers to the development of the toolboxes, which form the fundamental basis of the digitalization guide. The framework of the toolboxes is presented at the beginning and their characteristics are explained. The toolboxes are linked to the existing Industry 4.0 Maturity Index according to SCHUH ET AL. [25]. This index, postulated by SCHUH ET AL., describes different degrees of digitalization with the help of levels that build on one another and is already established in the industry. Finally, the procedure for developing the toolboxes is shown and explained.

2.1 Framework of toolboxes

In order to enable industrial companies to access Industry 4.0 in a transparent and comprehensive way, the entire production is divided into the following production areas: production planning and control and logistics, machine and plant engineering, quality management, and process technology (Fig. 1). Due to the fact that certain areas within battery cell production cannot be assigned to any specific area alone, an overarching area is added. For example, part of the overarching area relates to data. Since all production areas have to access it, no clear allocation can happen. The specific production areas in combination with the overarching area are the basis for the toolboxes. Each area is assigned its own toolboxes. All toolboxes provide content and methods that define development stages. These development stages represent different levels or stages of digitalization and Industry 4.0 and are based on the development stages from the Industry 4.0 Maturity Index according to SCHUH ET AL. [25]. The main terminologies used in this paper are summarized in figure 1.

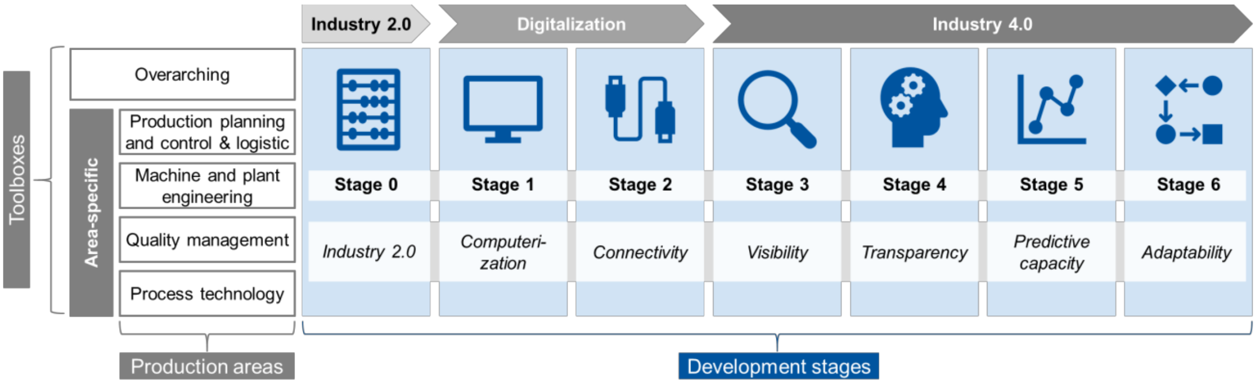


Figure 1: Overview of the framework toolboxes

The six stages represent the forms of digitalization and Industry 4.0 from computerization to adaptability. For the purpose of completeness, another stage 0 is added to represent the case of no digitalization. Stages 1 and 2 can be taken from the existing Maturity Index. Since the Maturity Index generally aims at the entire corporate structure, the definitions of levels 3 to 6 must be adapted to production. To achieve stage 3 sensors must be integrated into a process in order to capture product and process data. For example, in a calender, distance sensors could be integrated to continuously determine the existing layer thickness after the calendering process. In subsequent stage 4, the process-product-effect relationship is determined. At this stage, an interpretation is made of the influence of an input parameter on the output parameter. With the help of this understanding, a digital twin of the equipment or process can also be built and linked to the sensor data. In relation to the calendering process e.g. a correlation between roll gap and layer thickness could be determined. At the same time, the digital twin of the calendering process is built up and could visualize the effects of gap adjustment on layer thickness.

At the next higher level 5, the digital twin is used to predict the output parameters on the basis of the available input parameters. There is no direct link to the machine. In the case of calendering, the required roll gap can thus be predicted for a desired layer thickness, but only for a given material. The roll gap must then be set manually on the machine. The final stage adaptability is achieved when the system can not only predict the output, but also simultaneously adjust the system parameters accordingly. In the calendering process described above, manual adjustment of the roll gap is no longer necessary, since the system carries out the adjustment autonomously.

The framework of the toolboxes enables the user to classify his production quickly, comprehensively and clearly differentiated with regard to digitalization. Different areas can be targeted, as certain users or companies have different focal points. For example, a cell manufacturer might focus on the digitalization of its logistics, whereas a machine manufacturer of a calender might focus on process monitoring. Besides the stage analysis, a target stage to be achieved can also be defined, visualized and continuously tracked. In addition, the necessary steps to achieve the target stage can be taken from the toolbox. The toolboxes thus provide the decisive tool for a guided and comprehensive implementation of Industry 4.0 in battery cell production.

2.2 Toolbox development

The fundamental approach for developing each toolbox is shown in Fig. 2.

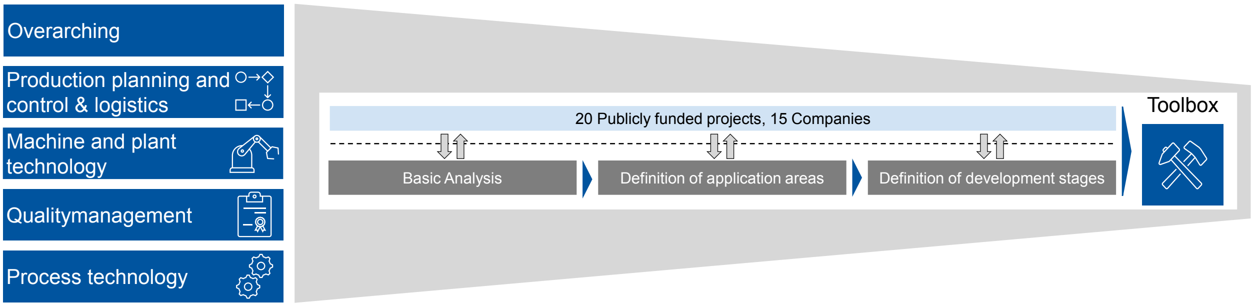


Figure 2: Procedure for the development of the toolboxes

The starting point for developing the toolboxes is a comprehensive basic analysis of the addressed production areas. To determine the status quo in battery cell production, the first step was to analyze pouch cell production by determining and summarizing all necessary input and output data along with established process and digitalization technologies. In addition, a project screening of 20 publicly-funded projects concerning battery cell production was carried out. Due to the currently running "Dachkonzept Forschungsfabrik Batterie" (meaning umbrella concept battery research factory) and various battery cluster initiatives of the Federal Ministry of Education and Research (BMBF) in Germany, there is intensive cooperation between all battery institutes operating in Germany. Based on the cooperation in the cluster "Intelligent Cell Production", there is close contact to parallel projects, which are regularly screened and surveyed within the framework of this project.

With the help of this project screening, the projects were classified according to the contents of the four production areas and their degree of digitalization. Furthermore, specific use cases of digitalization in battery cell production were developed from the individual projects. Alongside this screening, interviews were conducted with 15 industrial companies that currently support the BMBF Cluster Initiatives with a focus on battery cell production. The companies were e.g. cell manufacturers, equipment manufacturers, or specialists for sensor technology in the battery field. The interviews were conducted over a period of six months using a standardized questionnaire (multiple choice and open-ended questions) so that the results of all interviews could be compared.

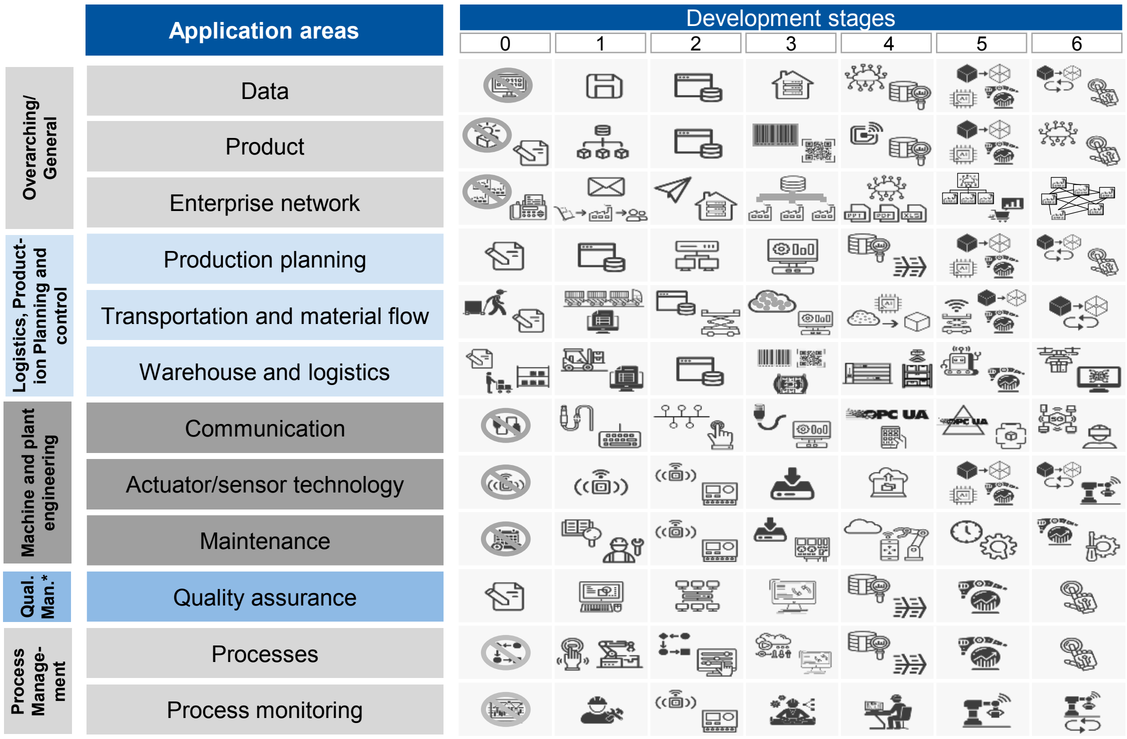
With the help of the interviews, it was thus possible to identify the current status quo of the implementation of Industry 4.0 and challenges in dealing with Industry 4.0. The main purpose of the interviews was to collect the various advances made by different players (material manufacturers, equipment manufacturers, cell manufacturers etc.) and to identify deficits. Based on the process analysis, the interviews, and the project screening, it is apparent that the specific production areas, but also the overarching area, need to be divided into different categories called "application areas". The application areas were determined based on a large number of interviews and the contents discussed. An attempt was made to aggregate and cluster the different topics. The application areas are not to be considered singular, as there are also interfaces between the different areas. For reasons of complexity, the interfaces are omitted at this point, so that they are not considered further for the time being.

The application areas are unique and customized according to the corresponding topic. Additionally, they describe applications within this area. Due to the application references, these are referred to as application stages in the following. In addition to the application area, the application stages form an area-specific basis for the integration of suitable Industry 4.0 methods. To build up the content of the development stages the identified use cases of digitalization in battery cell production are abstracted and used to assign specific Industry 4.0 content and methods. The resulting toolboxes of the different areas can be viewed in the further course as one large toolbox that provides all important information, content and methods.

3. Results and discussion

With the presented framework and procedure for the creation of the so-called toolboxes, 12 application areas were identified in total (see fig. 3). Multiple Industry 4.0 use cases from battery cell producers, plant manufacturers, and battery-related research projects were analyzed, clustered, and abstracted for general accessibility. The analysis concludes, that the toolboxes for the application areas can be divided into different sections. The first, general section can be applied to all production or company sectors.

For example, the application area "Data" can be assigned to the overarching toolbox since it accrues irrespective of the company's department. Data generation and analysis can occur at both the planning level and shop floor level. Other than the overarching application area, the following additional application areas can be identified: 1) Logistics, production planning and control, 2) Machine and plant engineering, 3) Quality and 4) Process management (see fig. 3). For each of these, toolboxes with defined characteristics based on the maturity index development stages are developed. These toolboxes, with their characteristics and methods, form the core element of a digitalization guide in battery cell production.



*Quality Management

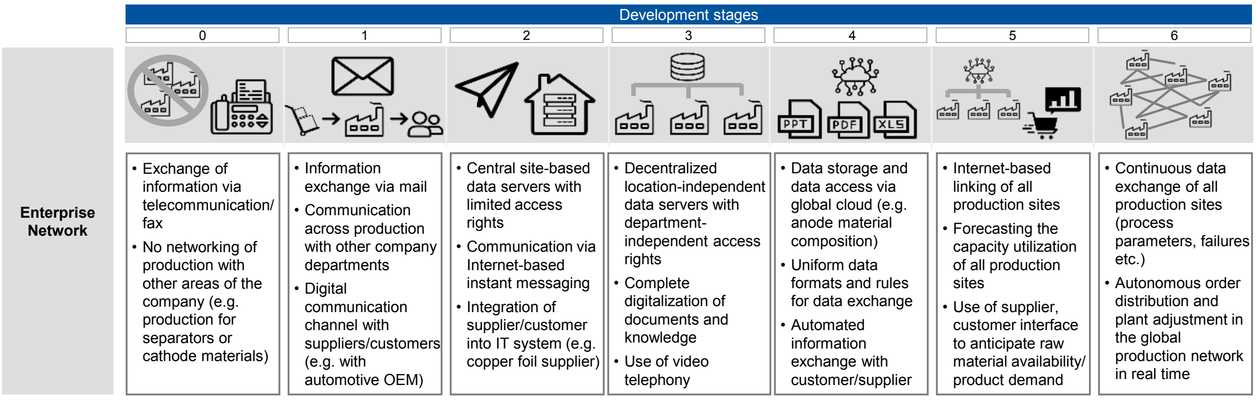
Figure 3: Application areas for digitalization in battery cell production

To identify or increase the level of maturity in regards to Industry 4.0, an analysis of the different toolboxes has to be conducted. Each application toolbox consists of relevant methods and tools to implement Industry 4.0 from a scale from zero to six in regards to the maturity index. To establish or increase the maturity level in the enterprise networks different approaches can be considered.

The application area “Enterprise Network” (see fig. 4) is presented in detail as an example toolbox and the process methodology to be used is shown. An enterprise network defines a company wide network as a form of coordinated cooperation and collaboration between several departments and locations. This toolbox focuses on approaches in the context of enterprise-wide networking. The starting point for the considerations is the question of how networking and collaboration can be optimized and costs reduced with the help of Industry 4.0. Improving the networking and communication of companies opens up synergies and avoids duplication of work. Networking production with other areas of the company, e.g. communication between the lithium-ion battery production facility and separator or anode material production facility results in unified IT solutions, standardized workflows, and consistent file formats that benefit the entire company. When improving the enterprise network up to the highest development stage six, whereas all data from the different production sites is continuously exchanged, an autonomous order distribution and plant adjustment in the global production in real time can be expected. To achieve this stage (six) of Industry 4.0, all previous stages from zero to five have to be achieved and implemented successfully in order to obtain this new development stage within the enterprise network.

The toolbox “Enterprise Network” provides detailed specifications of the development stages and therefore encourages company specific solution finding: the next development stages towards Industry 4.0 are always considered by the user(s) in the context of their company specific use case. This means, that the ideal or right target stage for the different production areas, can be an individual decision made by a company. This can result in different “ideal” target stages, depending on the current company orientation, but in the long term a high development level should be aimed for to ensure increased efficiency and quality in lithium-ion battery production.

In addition to the “Enterprise Network”, the other toolboxes enable companies to apply appropriate tools from the toolbox to their production in order to increase their level of maturity with regard to Industry 4.0. The other eleven toolboxes are structured and developed according to the same system and provide information on how to digitalize the warehouse, production processes, quality management, or how to implement predictive maintenance in production. The systematic and efficient development and implementation of digitalization, as well as the holistic assessment of a company's maturity with the toolboxes are enabled and provide an essential tool towards increased competitiveness. With the toolboxes, a more efficient Industry 4.0 development process can be achieved due to the strategic framework and the laid-out use case examples and best practices generated over the course of the interviews mentioned earlier. The existing application areas and development stages can be used and adapted according to the company's own needs, so that process development does not start from zero. The subsequent implementation of Industry 4.0 methods or applications in the company is also facilitated with the toolkits, as the contents of the toolkits build on each other and complement each other. The methodical overview can be a relief and increase in efficiency for an inexperienced user.



	Development stages						
	0	1	2	3	4	5	6
Enterprise Network	• Exchange of information via telecommunication/ fax • No networking of production with other areas of the company (e.g. production for separators or cathode materials)	• Information exchange via mail • Communication across production with other company departments • Digital communication channel with suppliers/customers (e.g. with automotive OEM)	• Central site-based data servers with limited access rights • Communication via Internet-based instant messaging • Integration of supplier/customer into IT system (e.g. copper foil supplier)	• Decentralized location-independent data servers with department-independent access rights • Complete digitalization of documents and knowledge • Use of video telephony	• Data storage and data access via global cloud (e.g. anode material composition) • Uniform data formats and rules for data exchange • Automated information exchange with customer/supplier	• Internet-based linking of all production sites • Forecasting the capacity utilization of all production sites • Use of supplier, customer interface to anticipate raw material availability/ product demand	• Continuous data exchange of all production sites (process parameters, failures etc.) • Autonomous order distribution and plant adjustment in the global production network in real time

Figure 4: Digitalization application and toolbox for industry 4.0 methods within an enterprise network

4. Conclusion and outlook

This paper presents a framework consisting of toolboxes for different application areas and stages for measuring and improving the maturity of Industry 4.0 in lithium-ion battery cell production. For this purpose, so-called toolboxes are used to determine the different stages of Industry 4.0 in different areas of battery cell production. Furthermore, actions that are needed to implement or increase Industry 4.0 in the different areas can be derived from the toolboxes. The main limitation and question remaining is concerning the right target stage for the different production areas, and how companies can define their target stages. These questions have to be analyzed and answered according to the companys’ specific strategy and mission. In summary, the toolboxes and the framework provide an overview and descriptions of tools or stages, but no guidance on how (e.g., how to implement or how to define a target state).

In the further course of the project, the methodology will be tested with the partners from lithium-ion battery industry and research. A workshop concept will be used for this purpose. To demonstrate the applicability of the methodology, the workshops will be validated in a broad field of industrial sectors in the area of battery cell production. In order for the toolboxes to be used, the terms Industry 4.0 and digitalization are defined at the beginning of each workshop. This is followed by an as-is analysis of the workshop participant using the toolboxes. Based on this, company-specific target states or wishes are defined, which signifies a higher development stage in the different application areas. When defining the target state, however, the cost-benefit ratio needs be considered and analysed by the individual company in question in order to ensure feasibility. The necessary analysis tools and methods will be developed in the next project phase as part of the overall methodology. Following the cost-benefit ratio, individual recommendations and/or projects are then derived from the target specification using creative methods. The defined measures or projects are prioritized and then implemented in the company. The aim of the validation workshops is to iteratively adapt the methodology to enable optimal measurability and further development of Industry 4.0 in battery cell production. The results will be summarised in a detailed guideline, which is intended to provide industrial companies with a transparent and holistic approach to Industry 4.0 in the field of battery cell production.

Acknowledgement

This paper is based on the research results of the project "Guideline for the digitalisation of battery cell production - BatterI4.0" [03XP0300A-D]. The project is funded by the German Federal Ministry of Education and Research within the framework of the competence cluster "Intelligent Battery Cell Production - InZePro". The authors of this publication thank the ministry for the funding.

References

[1] Michaelis, S., Rahimzei, E., Kampker, A., Heimes, H., Offermanns, C., Locke, M., Löbberding, H., Wennemar, S., Thielmann, A., Hettesheimer, T., Neef, C; Kwade, A., Haselrieder, W., Blömeke, S., Doose, S., Drachenfels, N. von, Drees, R., Fröhlich, A., Gottschalk, L., Hoffmann, L., Kouli, M., Leithoff, R., Rahlfs, S., Rickert, J., Schmidt, L. O., Schoo, A., Thomitzek, M., Turetskyy, A., Vysoudil, F. (2020). Roadmap Batterie-Produktionsmittel 2030 - Update 2020., VDMA Verlag GmbH, Frankfurt am Main

[2] Küpper, D., Kuhlmann, K., Wolf, S., Pieper, C., Xu, G., Ahmad, J. (2018) The Future of Battery Production for Electric Vehicles, Germany

[3] Saridogan, E., Güloglu, B., Hannum, C. (2020). The Effects of Technological Innovations on Competitiveness and Economic Growth, DOI: 10.26650/B/SS10.2020.001, Istanbul

[4] Li Da Xu, Eric L. Xu & Ling Li (2018) Industry 4.0: state of the art and future trends, International Journal of Production Research, 56:8, 2941-2962, DOI: 10.1080/00207543.2018.1444806.

[5] Dorrmann, L., Sann.Ferro, K., Heininger, P., Mähliß, J. (2021). Kompendium: Li-Ionen-Batterien Grundlagen, Merkmale, Gesetze und Normen. VDE Verband der Elektrotechnik Elektronik Informationstechnik e. V., Frankfurt am Main

[6] Riexinger, G., Doppler, J. P., Haar, C., Trierweiler, M., Buss, A., Schöbel, K., ... & Bauernhansl, T. (2020). Integration of traceability systems in battery production. Procedia CIRP, 93, 125-130.

[7] Huber, W. (2016). Industrie 4.0 in der Automobilproduktion. Springer Fachmedien Wiesbaden.

[8] https://northvolt.com/articles/connected-factory/, last accessed 14. March 2022, 14:03.

[9] https://www.freyrbattery.com/green-batteries/facility, last accessed 14. March 2022, 16:12.

[10] David R. Sjödin, Vinit Parida, Markus Leksell & Aleksandar Petrovic (2018) Smart Factory Implementation and Process Innovation, Research-Technology Management, 61:5, 22-31.

[11] https://www.catl.com/en/manufacture/, last accessed 14. March 2022, 18:07.

[12] Ayerbe, E., Berecibar, M., Clark, S. Franco, A., Ruhland, J. (2021). Digitalization of Battery Manufacturing: Current Status, Challenges, and Opportunities. Advanced Energy Materials 2102696

[13] Kiel, D., Müller, J. M., Arnold, C., & Voigt, K. (2017). Sustainable Industrial Value Creation: Benefits and challenges of industry 4.0. International Journal of Innovation Management, 21(08), 1740015.

[14] Thiede, S., Turetskyy, A., Kwade, A., Kara, S., & Herrmann, C. (2019). Data mining in battery production chains towards multi-criterial quality prediction. CIRP Annals, 68(1), 463-466.

[15] https://www.fraunhofer.de/en/press/research-news/2017/march/battery-production-goes-indust-rie-4-0.html, last access on 20. April 2020, 9:30.

[16] https://www.daimler.com/innovation/battery-factories.html, last accessed 20. April 2020, 9:34.

[17] https://www.daimler.com/innovation/case/connectivity/industrie-4-0.html, last accessed 20. April 2020, 9:42.

[18] https://press.siemens.com/global/en/news/battery-production-digital-twin-machine-builder-imple-ments-digitalization-strategy-solutions, last accessed 20. April 2020, 9:31.

[19] https://www.hannovermesse.de/en/news/news-articles/battery-manufacturing-relies-on-di-gitiza-tion-and-recycling, last accessed 20. April 2020, 9:34.

[20] Zäh, M., Vogl, W., Wünsch, G., & Munzert, U. (2004). Virtuelle Inbetriebnahme im Regelkreis des Fabriklebenszyklus. iwb Seminarberichte, 74.

[21] Westkämper, E. (2004). Fabrikplanung und-konfiguration mit Werkzeugen der digitalen Fabrik. ua; Technische Universität München/Institut für Werkzeugmaschinen und Betriebswissenschaften: Virtuelle Produktionssystemplanung: Virtuelle Inbetriebnahme und Digitale Fabrik. München

[22] Westkämper, E., Niemann, J., Warschat, J., Scheesr, A. W., & Thomas, O. (2009). Methoden der digi-talen Planung. In Handbuch Unternehmensorganisation (pp. 515-568). Springer, Berlin, Heidelberg.

[23] Turetskyy, A., Thiede, S., Thomitzek, M., von Drachenfels, N., Pape, T., & Herrmann, C. (2019). Toward Data-Driven Applications in Lithium-Ion Battery Cell Manufacturing. Energy Technology, 1900136.

[24] Fleischer, J., Bauer, J., Klee, B. & Spohrer, A. (2016). Efficient implementation of I4.0 with the VDMA toolbox based on use cases. Adaptive and Smart Manufacturing, 121-127.

[25] Schuh, G., Anderl, R., Gausemeier, J., Ten Hompel, M., & Wahlster, W. (Eds.). (2017). Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten. Herbert Utz Verlag.

Biography

Prof. Dr.-Ing. Achim Kampker (*1976) is head of the Chair of Production Engineering of E-Mobility Components (PEM) at RWTH Aachen University and Deputy Head of Fraunhofer Research Institution for Battery Cell Production (FFB). In addition to teaching at RWTH Aachen University, he devotes himself to numerous projects. These include his involvement in the "Agora Verkehrswende" and in the "Expertenrat Elektromobilität" of the NRW state government. In his career he has had great entrepreneurial success with multiple start-ups and spin-offs.

Prof. Dr.-Ing. Jürgen Fleischer (*1961) is the director of the Chair Machines, Equipment and Process Automation of the wbk Institute of Production Science at Karlsruhe Institute of Technology KIT and the spokesman of the BMBF competence cluster Intelligent Battery Cell Production InZePro. As a recognized member of the scientific community, he is member of various national and international advisory boards.

Prof. Dr-Ing. Klaus Dröder (*1968) is a university professor in the field of "Manufacturing Technologies and Process Automation" and head of the Institute for Machine Tools and Manufacturing Technology (IWF) at the Technische Universität Braunschweig. After holding several management positions in corporate research at Volkswagen AG, he played a key role in the founding and operational development of the Automotive Research Centre Niedersachsen, of which he was managing director until 2009. Today, in addition to being a university professor and institute director, he is also a member of the board of the NFF and the Open Hybrid LabFactory.

Prof. Dr.-Ing. Rüdiger Daub (*1979) has been head of the Chair of Production Technology and Energy Storage Systems at iwb at Technical University of Munich and at the Fraunhofer Institute for Foundry, Composites and Processing Technology IGCV since 2021. Previously Prof. Dr.-Ing. Rüdiger Daub has held several positions at BMW Group Munich, most recently as head of technology development and prototype construction for lithium-ion cells.

Jessica Schmied (*1995) is research assistant at the Chair of Production Engineering of E-Mobility Components (PEM) at RWTH Aachen University. She studied mechanical engineering with a specialization in production technology at RWTH Aachen University.

Alexander Puchta (*1996) is research associate at wbk Institute of Production Science at Karlsruhe Institute of Technology KIT. He graduated in mechanical engineering at the Karlsruhe Institute of Technology.

Timon Scharmann (*1995) works as a research associate at the Institute for Machine Tools and Production Technology at the Technische Universität Braunschweig as well as at the Battery LabFactory Braunschweig. He graduated in mechanical engineering at the Technische Universität Braunschweig with a specialization in production and systems engineering.

Hans-Christoph Töpper (*1994) graduated in mechanical engineering at the Karlsruhe Institute of Technology and is now working as a research associate at the TUMint. Energy Research GmbH and Institute for Machine Tools and Industrial Management (iwb) at the Technical University of Munich.

3rd Conference on Production Systems and Logistics

An Approach For Designing And Developing Microservices-enabled Manufacturing Operations Management In Greenfield Environments

Michael Oberle[1], Sascha Gärtner[1], Ozan Yesilyurt[1], Florian Maier[1], David Brandt[1]
[1]*Fraunhofer IPA / Stuttgart / Germany*

Abstract

Microservices are a special type of service-oriented architecture that promises flexibility through the encapsulation of functionality and loose coupling. With a growing need for responsiveness of manufacturing systems driven by customer expectation, competition, and regulatory changes, the microservices architecture is highly attractive for Manufacturing Operations Management (MOM) solutions. Manufacturing enterprises that have a traditional MOM solution as a monolithic system in place can transition to microservices architecture through reverse engineering their existing solution. This is not possible in a greenfield environment where the entire manufacturing system is built from scratch. In this paper, we present an approach for designing and developing a microservices-enabled MOM solution for greenfield environments. The approach is based on industry standards and incorporates a parallel design and development of the microservices addressing the functional needs of the MOM solution and equipment emulators, enabling early solution testing before the actual equipment is delivered. We evaluate the approach through a case study in which we developed a microservices-enabled MOM solution for a small-scale battery manufacturing line. We report that the ISA95 standard provides a good guideline for microservice MOM solutions, though the underlying data model needed to be modified to cover our case. The developed instantiation also shows that a microservices-enabled MOM can flexibly address new requirements. However, this does not come without a cost, as their integration and management require additional effort compared to monolithic architectures.

Keywords

Manufacturing Operations Management; Microservices; Manufacturing Execution System, MES; MOM; IIoT; Battery Cell Manufacturing

1. Introduction

The need for flexibility of manufacturing companies to adapt their operations to changes in the environment has become apparent in current crises such as COVID19 or the Ukraine crisis, in which the supply chains have been heavily disrupted. However, if even these crises would not exist, manufacturers need to be responsive to adapt to external changes [1]: Customers expect new product releases in a shorter timeframe, there is increasing global manufacturing competition, and governments impose new regulations to act on climate change. To deal with the challenges manufacturing companies must embrace flexibility throughout their manufacturing system, from the material over machinery to IT-Systems. With the increasing importance of data in manufacturing processes, the flexibility of the manufacturing IT-Systems that support the management of manufacturing operations (MOM) is crucial [2,3]. MOM solutions support through their underlying model a certain degree of flexibility to address changes in the manufacturing system, such as the addition of new products or production steps [4]. However, not all changes that occur later in practice can be foreseen. To dissolve this limitation, the adoption of microservices architecture, in which the MOM

DOI: https://doi.org/10.15488/12132

solution consists of several microservices, was proposed [5]. Microservices architectures will not necessarily avoid source code changes, but they should at least reduce these changes due to loose coupling. This capability makes a microservices-enabled MOM highly interesting for both established and new manufacturing sites. However, the methodology for designing and developing such a system substantially differs. Established manufacturers can reverse engineer their existing MOM solutions and gradually transition to the new architecture paradigm [6] [7]. New manufacturing sites are planned from scratch, and thus guidance for the MOM design and development endeavour is required. This paper presents an approach for designing and developing a microservices-enabled MOM that we evaluate by implementing a MOM solution in a case study. The remainder of the paper is structured as follows. In section 2, we discuss related work on MOM solutions and highlight the gap in research. Section 3 presents our approach for designing and developing a microservice MOM solution. We apply this approach in section 4 in a case study for a new small-scale battery manufacturing line. We evaluate our approach and our MOM solution in section 5. Finally, we conclude our paper with the key points in section 6.

2. Related work

For this section, we conduct a literature review on the design of MOM solutions. We searched Google Scholar with a focus on the last five years and extended our retrievals through cross-reference search. Among the literature, we identified three common themes in our review: the importance of standards for MOM solution development, the requirements of MES in the context of Industry 4.0, and the potential of microservices for increasing the flexibility of MOM solutions. Considering standards for MOM development, ISA95 is a well-known standard. [8] uses ISA95 as a guideline for developing an MES based on the open-source enterprise software platform odoo. [9] leverage the data model of ISA95 to implement an IIoT connected MOM system. Interestingly the model seems not to be applied in practice for system development, at least in some industries. A comparative study of different MES vendors [10] reports that only very few systems comply with the ISA95 standard. Other studies highlight integrating MES with IIoT platforms to align with Industry 4.0. [11] review relevant standards and ontologies of MES in the context of Industry 4.0. They highlight that an MES needs to be seamlessly integrated with all cyber-physical system components to enable highly automated solutions. Formal ontologies and models play an essential role in ensuring interoperability. [9] combine a quality function deployment and case studies to identify MOM's main requirements in smart factories. They also identify interoperability as the key requirement for MOM in a smart factory. This is also supported by [12], who identify interconnectedness as a crucial feature of system architectures for Industry 4.0. Several authors point out modularity as a desirable characteristic of next-generation MOM solutions. In particular, the microservices architecture is suggested as a suitable solution to address modularity [9,12, 5]. [13] point out that the functional scope of the service has an impact on the communication infrastructure and should be considered in the design of a solution but also argues that the benefits of flexibility are highly suitable for smart manufacturing. Aligned with this claim is a slowly growing interest in MOM solutions based on microservices. There is a small number of implementation attempts. These attempts are compared in table 1. Our literature review analysed the available approaches concerning the other two common themes. The available implementations do not follow any standard for their implementation but instead, seem to be built from intuition. In addition, with exception of [14], the connectedness to other systems as part of a smart factory was not considered. We further identify a gap in a missing methodology or approach for designing and developing a MOM solution in a greenfield environment. In our work, we aim to address this gap.

Table 1: Comparison with related work on MOM microservice architectures

Table	Li et al. 2019	Zhou et al. 2019	Wunck & Jonas 2019	Jin et al. 2021	Yi et al. 2019

Target	Greenfield	Greenfield	Brownfield	Conceptual	Conceptual
Model	Not specified	Not specified	Not specified	Not specified	Not specified
IIoT Integration	Yes	No	No	No	No
Case Domain	Production network	Not specified	Laser-Cutting	Tobacco manufacturing	None
Case Environment	Demonstrator	Not specified	CPS testbed	Conceptual	None

3. Approach for designing microservice-enabled MOM in a greenfield environment

This section presents our approach to designing a microservice MOM solution in a greenfield environment. The approach is visualised in Figure 1 including several steps and the associated vital roles involved. In contrast to the brownfield design of MOM solutions in which the design can orient on the existing boundaries, such as the equipment and software in place, the design for a greenfield does not have such strict constraints in place [15]. However, it is also evident that there must be a goal in place in which the principal purpose and scope of the manufacturing line are defined. There is a substantial difference in the design of the entire manufacturing line if the purpose of the manufacturing line is the production of high-volume goods compared to a manufacturing line for a small number of individualised products. For high-volume production, throughput is the key, and by contrast, flexibility is vital for the production of individualised products. The purpose impacts the entire manufacturing system, from equipment over processes to the MOM solution. For this reason, the first step in our approach is the definition of the manufacturing line goal and scope (0). This is the task of the executives overseeing the entire manufacturing site project. This definition serves as an input for the following steps concerned with the MOM solution's design. Our approach includes a parallel line to the steps for the actual MOM solution, which are targeted at creating equipment emulators. This is necessary in a greenfield environment where equipment is typically not present from the beginning. This defers gaining experience with the MOM solution, which could lead to severe flaws in the design. These are harder to fix the latter they are discovered [16]. The line for MOM solution design starts with the definition of operational procedures (1a). This step defines how operators on the shop floor interact with a MOM solution and the equipment of the shop floor. ISA95 part 1 defines a function model of manufacturing enterprises. Based on the defined goal and scope definition, the relevant functions can be selected and guide the characterization and modelling of the procedures. The result serves as the input for the segmentation of microservices of the MOM solution (2a). In this step, the functionalities of the microservices are determined, which defines the scope of each microservice. ISA95 part 3 provides activity models for operations management, production, maintenance, quality, and inventory. Each model activity is well-defined with clear boundaries and relationships to other activities. This property renders these models a viable template for segmenting a MOM solution in microservices. Each MOM microservice should have a clear functional scope [5]. Therefore, activities in the operational definition need to be mapped on these activity models. In the following step, the detailed design of the individual microservices takes place (3a). This involves designing the underlying semantics that determine the interactions of the microservices. ISA95 part 2 defines several common object models for all major entities involved on the shop floor, including personal, material, and equipment on which the data model design can be based. The parallel line of design of the equipment emulators begins with the manufacturing process characteristics (1b). The VDI 3682 standard provides a concept for describing processes in a standardized format. The format is solution neutral and allows the description of a process as the desired behaviour of a system by inputs and outputs in terms of product, energy, and information of the system. The latter is the most important for the emulator design.

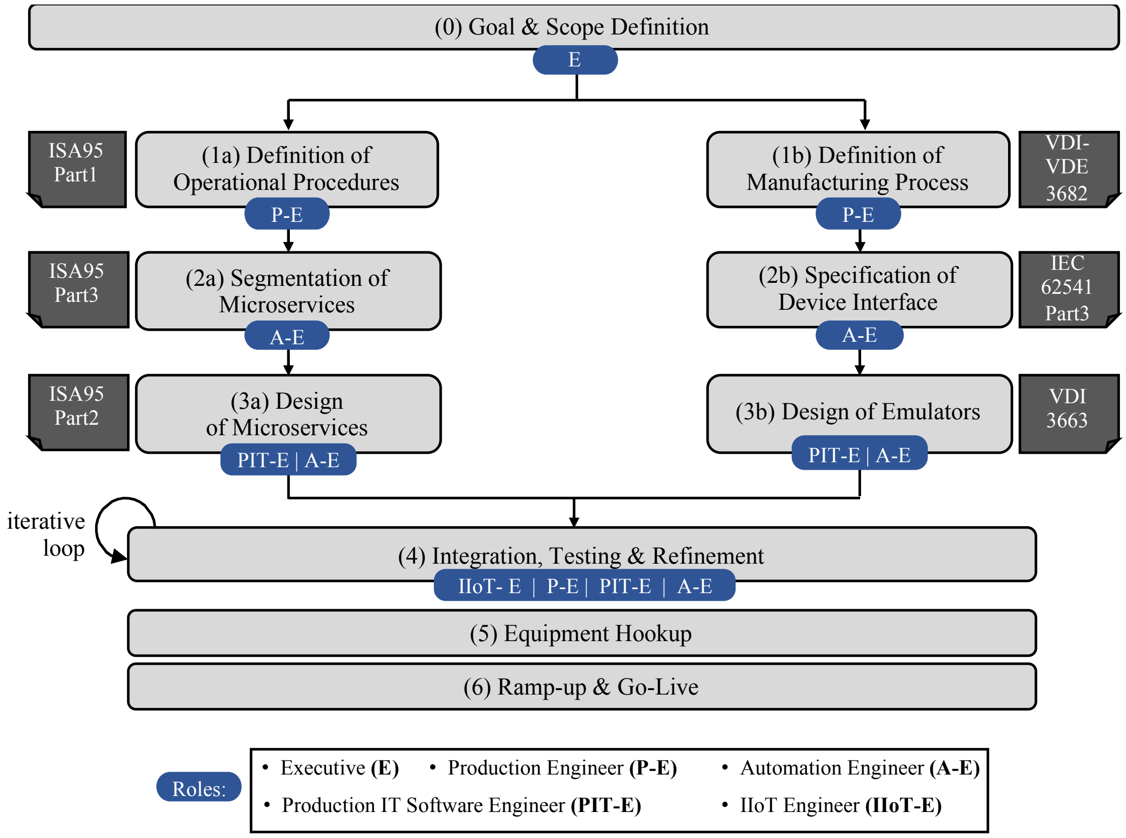


Figure 1: Approach for designing and developing a microservices-enabled MOM in a greenfield environment

Production engineers who design the manufacturing good's physical transformation process are the source of this information. Based on this specification, the information interface of the equipment can be defined in the following (2b). OPC-UA has been becoming a de-facto standard for the interoperability of equipment with MOM solutions in smart factories [17]. The IEC62541 standard specifies the protocol and provides the elements for constructing the information model from an IT perspective. These include different nodes such as objects, variables, and methods. An Automation Engineer needs to structure the IT interface based on the information inputs and outputs from the formal process description of the previous step. The emulator's behaviour must be specified in the last step of the equipment emulator design (3c). The behaviour model can be implemented through simulation. VDI 3663 describes guidelines for the construction simulation models for material flow and production systems. Based on this standard and process, the automation engineer must specify the behaviour of the underlying model. After completing the design for the MOM microservices and the equipment emulators, the development begins [4]. Given the distributed nature and required interaction of microservices architectures [5] we propose an iterative approach that aims at an early end-to-end integration of all solution components to identify defects in design or implementation as soon as possible. State of the art microservices architectures are hosted on IoT platforms. For this reason, the development requires the collaboration between Production IT Software Engineers who possess the expertise to implement the functionality and IIoT Engineers who provide the know-how of integration and deploying the software. Automation and Production Engineers need to be involved in testing to provide recommendations for improvements to the MOM solution. The development iterations continue with the equipment emulators as long as the actual equipment is unavailable. As the implementation of the equipment interface might be itself subject to development defects on the side of the equipment supplier, it is essential to replace the equipment emulator during testing as soon as possible with the actual item of equipment (5). Testing with the actual equipment allows the development of a mature MOM solution that can be used early in the ramp-up to accelerate the final go-live of the manufacturing line (6).

4. Case Study: Small scale battery manufacturing line

Aligned with the Design Science Research Methodology [18] we applied the previously presented approach in a case study for designing a real microservices-enabled MOM solution to evaluate its utility. The battery manufacturing pilot line within the Center for Battery Manufacturing (ZDB) at Fraunhofer IPA served as the application environment. In 2019 the project for establishing a small-scale manufacturing line was initiated with the goal to produce small quantities of round li-ion cells, enabling research on new battery materials and manufacturing technologies. The entire manufacturing line was planned to cover all battery manufacturing steps, from slurry mixing over electrode coating to cell assembly and formation. The project focused on establishing the battery cell process steps from scratch in the past two years. This involved provisioning equipment for ten sub-processes. In the following, we report on our experience applying the previously described approach for designing and developing a microservice-enabled MOM in this specific greenfield case and present our developed artefact.

4.1 Application of the approach

For our MOM solution endeavour, the definition of goal and scope (step 0) was given through the ZDB project's previously described goal. Our work started with defining the operational procedures and the battery cell assembly processes that should be part of ZDB. We conducted these steps in collaboration with production engineers responsible for the battery cell assembly. To define the scope of the operational procedures, we analysed the functional model of ISA95 part 1 (step 1a). This includes ten functions for enterprise control of a manufacturing organization. Considering the goal of our manufacturing line, we identified that only order processing, production scheduling, production control, quality assurance, and product inventory control were relevant for our project. These are the core functions that are required to produce battery cells. At this stage and scale of the ZDB, the other functions are not (yet) relevant as the ZDB does not aim at manufacturing goods for commercial purposes. These functions themed the frame for the procedures we needed to consider in the operational procedure definition. The Production Engineers defined ten sub-processes for the battery cell assembly (Figure 2).

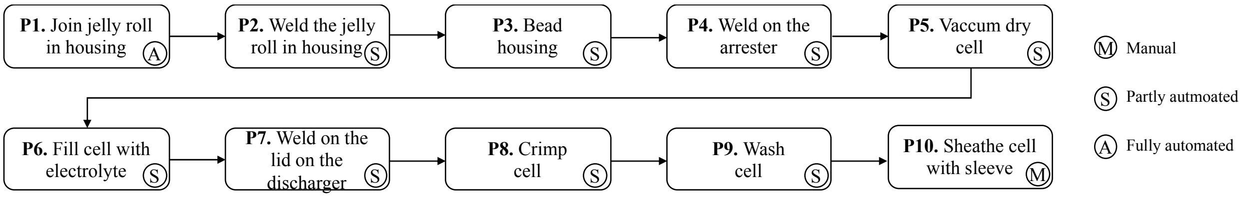


Figure 2: Battery Cell Assembly sub-processes

We jointly modelled the details of the planned operating procedures with UML activity diagrams for each of these sub-processes. These include swimlanes for the human actor (engineer or worker), the manufacturing equipment, and the MOM solution. For each item of equipment in the sub-processes P1 to P10, we conducted a specification of the process information inputs and outputs (step 1b), including units and measurement frequency. We sub-categorized thereby outputs to information available at the end of each sub-process, such as the result of the process or a particular measurement (e.g. height of groove) and information that characterizes the physical process during its execution (e.g. forces or temperatures). Based on the results from step 1a, we started the segmentation of the microservices (step 2a) by mapping the activities of the MOM solution in the UML activity diagrams to the activity model of production operations management in ISA95 part 3. Except for the detailed production scheduling activity and product resource management, which we decided should not be covered by the MOM solution due to the small production quantities, all activities were relevant. The activities related to production dispatching, product definition management, and production data collection determined the scope of single microservices. Due to the relatedness of production tracking and production execution, we decided to combine these activities in a single microservice. We also identified that most subprocesses (see Figure 2) required manual interaction of the worker. Therefore, there needed to be a way to convey recipe information. For this reason, we introduced another microservice, the operator guide that should cover this functionality. We specified the equipment interfaces (step 2b) using the results from step 1b as input. We used essential elements from the IEC62541 to construct a lightweight information model in cooperation with the Automation Engineer. This lightweight information model consisted of four central nodes: machine state, machine input, machine output, and process measurements. The latter directly reflects the categories of the process specification in step 1b. We introduced the respective variables in each corresponding node. For the detailed design of the microservices, we analyzed the production operations information model of ISA95 part 2 with a focus on the requirements of our case. We identified that the model itself has the most relevant entities, such as product definitions. However, it does not provide the flexibility for deriving product definition quickly from existing variants. Therefore we designed a new semantic model as a development foundation (see Figure 3).

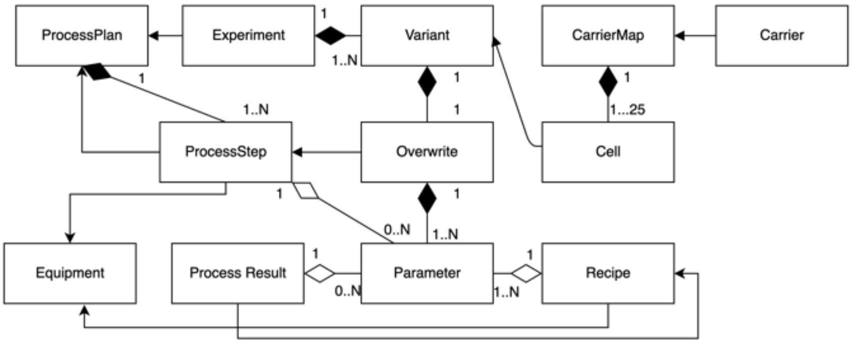


Figure 3: Semantic Data Model

We concluded the detailed design of the emulators (step 3b) by defining the behaviour of the equipment emulators. The VDI 3663 standard describes a detailed description of how simulations can be developed. Due to cost-benefit reasons, we decided against implementing a detailed physical simulation. Instead, we defined a range of reasonable values for each of the variables in the OPC-UA interface and a distribution function for statistically controlling value occurrences. We started the development activities by identifying artefacts that we already had available from previous projects and could be re-used to develop the MOM solution. These were the IIoT platform Virtual Fort Knox [19] including the integration middleware Manufacturing Service Bus (MSB) [20], a platform for creating emulators with OPC-UA and a lightweight MES from a partner organization. The MES could not cover all the required functionality for the defined operational procedures, particularly the management of experiments for single cells. However, it provides an API for controlling the functionalities for tracking and executing products through a manufacturing process which is why we used it as the solution for this microservice. We implemented the equipment emulators by transferring the equipment IT interface specification to an OPC-UA model using the SiOME modelling editor. The formal specification is the basis for instantiating an instance in the equipment emulation platform. The behaviour for each node was then specified by a separate configuration that includes the range, the distribution function, and the refresh rate. For the implementation of the MOM solution, we used an agile methodology defining a product backlog. We prioritized the implementation of a stub for each microservice that was able to be deployed on VFK. These stubs could communicate the most relevant information of the entities described in the semantic model via the MSB and were able to interact fundamentally with the equipment emulators. All microservices were further refined based on the feedback from the Production Engineers. As soon as an item of equipment arrived, we verified the integration of the equipment with the MOM solution. Apart from minor fixes, such as the OPC-UA security levels, we did not encounter issues. As one of the last steps, we integrated the workpiece carrier, for which we conducted three development sprints until no problems could be identified any longer. The final go-live of the battery cell assembly line is still pending at the date of writing due to the safety certification of the line, which was affected by the COVID pandemic.

4.2 Instantiation of the microservice-enabled MOM solution

The instantiation that resulted from applying the proposed approach is presented in Figure 4.

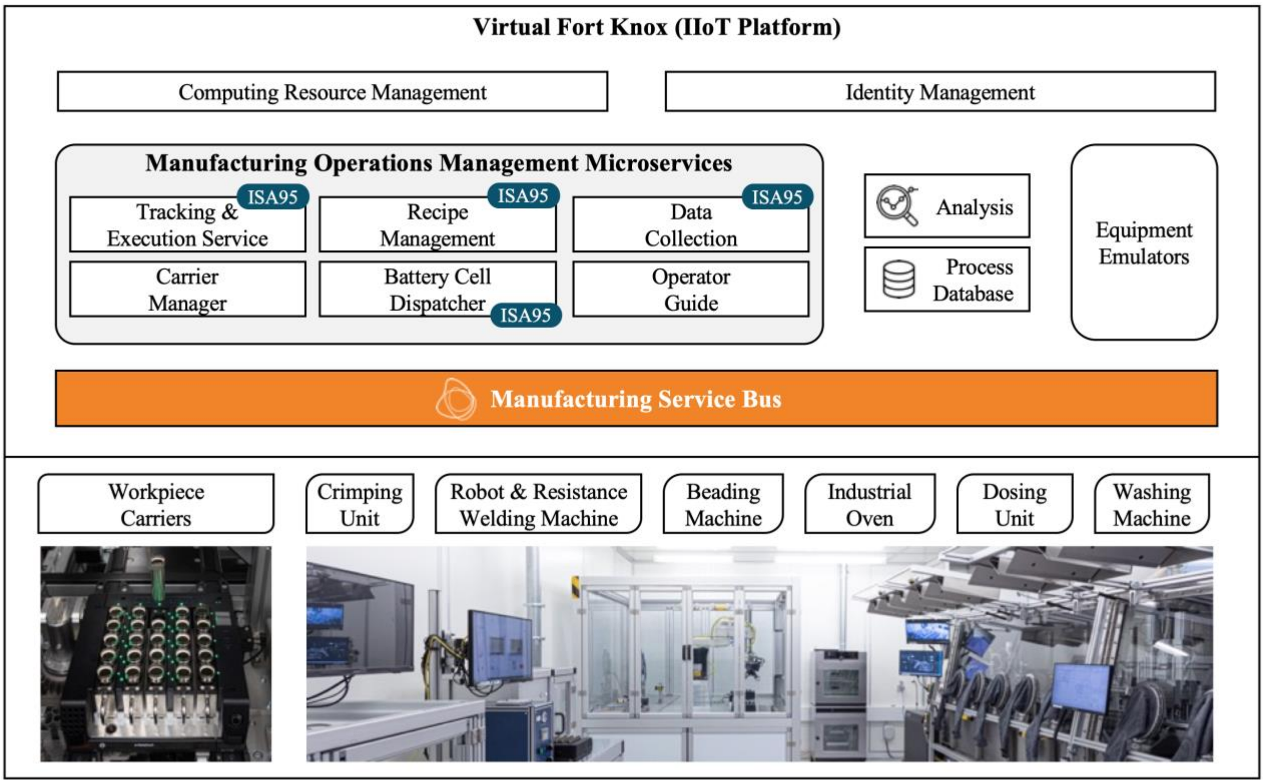


Figure 4: Overview of the MOM architecture

The VFK platform provides the provisioning of computing in the form of virtual machines. In our case, we include an additional layer of virtualization by embodying each artefact in a separate docker container to simplify the deployment. In addition, the platform provides the basic service of managing users' accounts for all microservices (identity management) and the MSB enabling the integration of microservices and shop floor equipment. The MSB also provides important security features. Only communication of registered and authorized microservices and equipment that passes over a secured connection channel is allowed. MOM is supported in our architecture through six microservices. The tracking and execution microservice allows the management of the process plans for battery cells. It tracks the progress of the cells through the individual manufacturing steps. The recipe management enables the definition of parameter variants for all individual process steps. Existing experiment setups can be used as templates. The battery cell dispatcher decides which cells are processed next. This is important as single cells can be extracted manually from the workpiece carrier. The carrier manager enables the placement of individual cells on the workpiece carrier. It allows the assignment of recipes to the single cells in the carrier. The operator guide conveys recipe information for processing the following cells to the worker. In addition, it allows the indication of process stops and starts. Data collection records all measurements during process execution (e.g. forces) with the context of the individual cell and recipe parameters. Our overall architecture also includes data storage and process analysis services, which are realized through pre-configured virtual machines for Grafana and InfluxDB provided by the VFK platform. Figure 5 presents the basic interaction sequence of our microservices-enabled MOM.

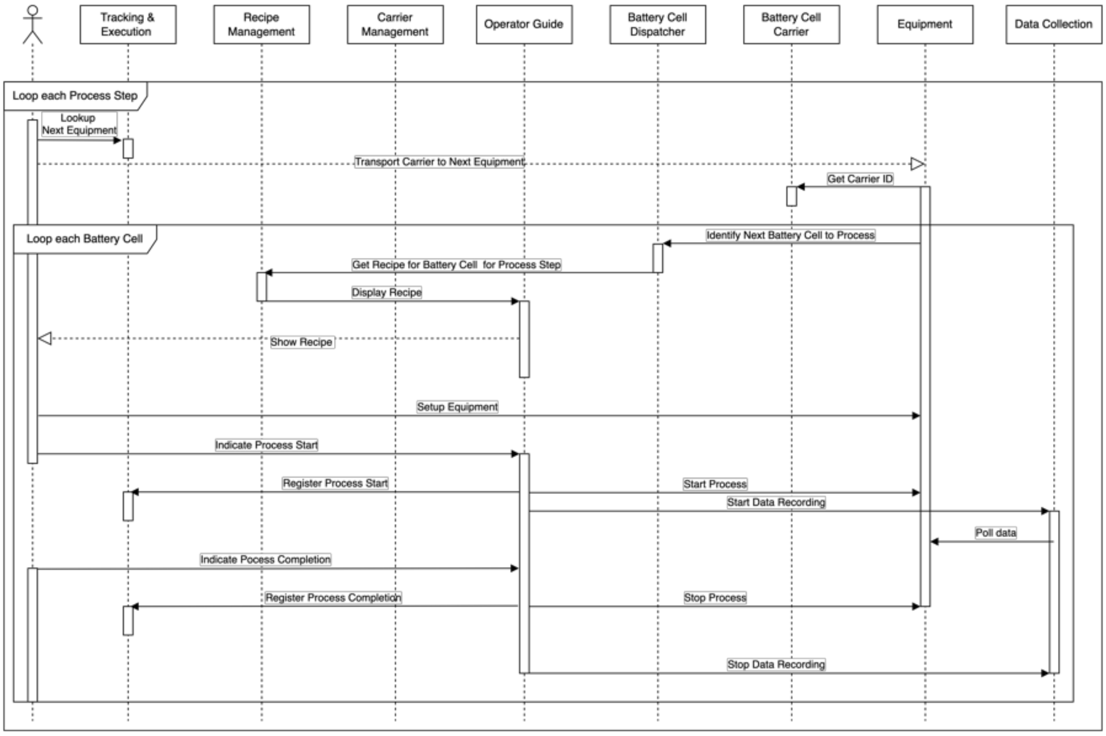


Figure 5: Interaction of microservices for experiment execution

5. Evaluation

In the following, we evaluate our approach and our instantiation by reflecting on our design and development experiences.

5.1 Approach

As our approach leads to an instantiation that, according to the feedback from the tests of our Production Engineers are covering the procedures adequately, we conclude that our approach can deliver effective MOM solutions on microservices. We see the merit of our approach on the one hand in the practical guidance by the incorporation of standards. ISA95 provided good advice for eliciting the functional requirements. On the other hand, the parallel use of emulators proved very valuable. Due to the global demand for battery manufacturing equipment, the delivery of our equipment was severely delayed. However, our emulators allowed us to verify our solution early. This showed during the integration of the workpiece carrier, for which we did not have an emulator but only interface specifications. In contrast to the actual manufacturing equipment in which only minor fixes were necessary, we needed three full days to integrate the workpiece carrier without any remaining issues. However, we also see the potential for improving our approach. The specification of the interface could benefit from OPC-UA companion specifications [21]. This would streamline the interface definition process and reduce the probability of issues even further.

5.2 Instantiation

Although the battery cell assembly line is not yet operational, we can already report from our experience regarding a necessary change for the MOM solution. In our design of the carrier management and the cell dispatcher microservice, the specification that we worked on with the Production Engineers only foresaw those stations would process the entire workpiece carrier as a whole or single cell. However, later during the project, it became clear that the workpiece carrier could not be placed in harsh environments like the washing machine or the oven. For this reason, trays on the workpiece carrier were introduced. The necessary change could be implemented by the carrier manager and cell dispatcher without affecting any other services. With a monolithic solution, this would have had a high probability of breaking changes affecting other modules. However, we also experienced drawbacks from the microservices architecture. Testing is far simpler for monolithic architectures as the test setup does not have to spawn multiple hosts and technologies, as it is the case for a microservices architecture. Another related issue is changes that might be necessary due to security vulnerabilities. When the log4j flaw [22] was discovered earlier this year, this meant investigating all microservices and their dependencies. In a monolithic architecture, such investigations are easier to handle.

6. Conclusion

Promising increased flexibility, microservice architectures are an attractive architecture style for MOM solutions. There are some attempts at implementing microservices production management solutions. Covering the lack of guidance for designing and developing such solutions for greenfield environments, we presented an approach building on industry standards. The approach incorporates a parallel design of equipment emulators and MOM microservices to be able to verify the solution as soon as possible, even when the equipment is not yet delivered. The ISA95 provides a valuable baseline for defining the operating procedures necessary for eliciting the functional requirements and segmenting the microservices. In a case study on a MOM solution for battery cell assembly, we could successfully develop a solution that covers the needs of the Production Engineers. Our experience with our instantiation indicates that flexibility can be achieved through a microservices-enabled MOM solution. However, there are also costs in managing complexity and testing that need to be addressed in future research.

Acknowledgments

This work was partially supported by a grant from Federal Ministry of Education and Research providing for this work through the project "DigiBattPro 4.0" and the State Ministry of Baden-Wuerttemberg for Economic Affairs, Labour and Tourism (Project »Zentrum der Digitalisierten Batteriefertigung«).

References

[1] Greeven MJ, Howard Y, Shan J. Why Companies Must Embrace Microservices and Modular ThinkingGreeven, Mark J., Howard Yu, and Jialu Shan. MIT Sloan Management Review. 2021;62(4):1–6.

[2] Mantravadi S, Møller C. An Overview of Next-generation Manufacturing Execution Systems: How important is MES for Industry 4.0? Procedia Manuf. 2019;30:588–95.

[3] Almada-Lobo F. The Industry 4.0 revolution and the future of Manufacturing Execution Systems (MES). J Innovation Management. 2016;3(4):16–21.

[4] Mantravadi S, Srai JS, Brunoe TD, Møller C. Exploring Reconfigurability in Manufacturing Through IIoT Connected MES/MOM. 2020 Ieee Int Conf Industrial Eng Eng Management Ieem. 2020;00:161–5.

[5] Wunck C. Towards a Microservice Architecture for the Manufacturing Operations Layer. In: Proceedings of 32nd International Conference on Computer Applications in Industry and Engineering. 2019. p. 241–230.

[6] Francesco PD, Lago P, Malavolta I. Migrating Towards Microservice Architectures: An Industrial Survey. 2018 Ieee Int Conf Softw Archit Icsa. 2018;29–38.

[7] Lauretis LD. From Monolithic Architecture to Microservices Architecture. 2019 Ieee Int Symposium Softw Reliab Eng Work Issrew. 2019;00:93–6.

[8] Li C, Mantravadi S, Møller C. AAU Open Source MES Architecture for Smart Factories – Exploiting ISA 95. 2020 Ieee 18th Int Conf Industrial Informatics Indin. 2020;1:369–73.

[9] Mantravadi S, Møller C, LI C, Schnyder R. Design choices for next-generation IIoT-connected MES/MOM: An empirical study on smart factories. Robot Cim-int Manuf. 2022;73:102225.

[10] Kannan SM, Suri K, Cadavid J, Barosan I, Brand M van den, Alferez M, et al. Towards Industry 4.0: Gap Analysis between Current Automotive MES and Industry Standards Using Model-Based Requirement Engineering. 2017 Ieee Int Conf Softw Archit Work Icsaw. 2017;29–35.

[11] Jaskó S, Skrop A, Holczinger T, Chován T, Abonyi J. Development of manufacturing execution systems in accordance with Industry 4.0 requirements: A review of standard- and ontology-based methodologies and tools. Comput Ind. 2020;123:103300.

[12] Azarmipour M, Elfaham H, Gries C, Kleinert T, Epple U. A Service-based Architecture for the Interaction of Control and MES Systems in Industry 4.0 Environment. 2020 Ieee 18th Int Conf Industrial Informatics Indin. 2020;1:217–22.

[13] Götz B, Schel D, Bauer D, Henkel C, Einberger P, Bauernhansl T. Challenges of Production Microservices. Proc Cirp. 2018;67:167–72.

[14] Li P, Jiang P, Liu J. Mini-MES: A Microservices-Based Apps System for Data Interconnecting and Production Controlling in Decentralized Manufacturing. Appl Sci. 2019;9(18):3675.

[15] Rinderle-Ma S, Mangler J. Business Process Management, 19th International Conference, BPM 2021, Rome, Italy, September 06–10, 2021, Proceedings. Lect Notes Comput Sc. 2021;3–14.

[16] Hibino H, Fukuda Y. Emulation in manufacturing engineering processes. 2008 Winter Simul Conf. 2008;1785–93.

[17] Kannoth S, Schnicke F, Antonino PO. Enabling Industry 4.0 Communication Protocol Interoperability: An OPC UA Case Study. 7th Conf Eng Comput Based Syst. 2021;1–9.

[18] Hevner AR, Corwin EJ, March ST, Bozoky I, Park J, Pugh LC, et al. Design science in information systems research. MIS Quarterly: Management Information Systems [Internet]. 2004 Mar 1;28(1):75–105. Available from: http://dl.acm.org/citation.cfm?id=2017217

[19] Holtewert P, Seidelmann J, Wutzke R, Seidelmann J, Bauernhansl T. Virtual Fort Knox Federative, Secure and Cloud-based Platform for Manufacturing. Procedia CIRP [Internet]. 2013;7(0):527–32. Available from: http://www.sciencedirect.com/science/article/pii/S2212827113002965

[20] Schel D, Henkel C, Stock D, Meyer O, Rauhöft G, Einberger P, et al. Manufacturing Service Bus: An Implementation. Proc Cirp. 2018;67:179–84.

[21] Stoop F, Ely G, Menna R, Charache G, Gittler T, Wegener K. Smart factory equipment integration through standardised OPC UA communication with companion specifications and equipment specific information models. Int J Mechatronics Manuf Syst. 2019;12(3–4):344–64.

[22] Korn J. The Log4j security flaw could impact the entire internet. Here's what you should know - CNN [Internet]. 2021 [cited 2022 Mar 25]. Available from: https://edition.cnn.com/2021/12/15/tech/log4j-vulnerability/index.html

Biography

Michael Oberle (*1984) is a computer scientist and group leader at Fraunhofer IPA working in the field of smart manufacturing with a focus of data-driven and event-driven production control services.

Sascha Gärtner (*1993), is a research associate at Fraunhofer IPA working on the practical application of artificial intelligence methods in the smart manufacturing environment.

Ozan Yesilyurt (*1991), is a research associate at Fraunhofer IPA researching IT system landscapes for battery manufacturing and energy flexibility with AGV batteries in manufacturing companies.

Florian Maier (*1983), is a research associate at Fraunhofer IPA researching resilience strategies for manufacturing companies and manufacturing operations management solutions.

David Brandt (*1992), is a technical employee at the Fraunhofer IPA supporting the team with his expertise in database and web development IT systems design and development projects.

3rd Conference on Production Systems and Logistics

Methodical Approach For Detailed Planning Of Services To Offer Product Service Systems

Dominik Lins[1], Bernd Kuhlenkötter[1]
[1]Ruhr-University Bochum, Chair of Production Systems, Bochum, Germany

Abstract

The transformation of current business models towards offering product service systems (PSS) provides manufacturing companies numerous opportunities to consolidate or even expand their competitive position. Companies are confronted with the challenge of successfully designing this transformation process simultaneously. In order to approach the development of new business models and the transformation process, business model patterns and best practices provide a good first orientation for companies. However, these are designed to be industry-neutral and rather abstract when considering the actual processes. Thus, they do not offer any individual support to companies in the specific development of a business model and its required service delivery processes. Service delivery processes are part of a business model and describe activities that take place to provide services. Small and medium-sized enterprises (SMEs) in particular do not have the necessary time, technical and methodological resources to manage a transfer from abstract business model examples to an individual business model. This barrier often leads SME to remain with their traditional business model.

Therefore, this paper presents a methodology for the detailed planning of service delivery processes. The designed methodology supports the phases design and implementation, which are part of the business model development. The methodology describes a structured procedure, in which relevant services first have to be identified. These services are then broken down into individual process modules on a second level. The modules are elements that can get combined to services. On a third level there are explicit process models. The process models are assigned to the modules and define the respective process steps and the requirements for the implementation. The approach is designed to support companies successfully transform to new business models for PSS by applying the detailed planning for services with specific modules that contain detailed process models and requirements.

Keywords

product service systems; business model transformation; process modeling; detailed planning; SME

1. Introduction

Manufacturing companies, especially small and medium-sized enterprises (SMEs), face the challenge of consolidating or even expanding their competitive position in the increasing global market [1]. Ongoing digitalization and the development towards industry 4.0 drives and enables the transformation of current business models towards offering product service systems (PSS) and at the same time help companies to cope with the mentioned challenges [2,3]. A PSS is an integrated offer of one or more goods and services [4]. Companies are therefore increasingly concerned with the development of innovative business models

DOI: https://doi.org/10.15488/12187

for PSS that include the actual product combined with integrated services as individual product solutions. This creates greater added value for customers through more customized products, while companies also benefit through, for example, closer customer ties and continuous contact during the service delivery phase [5]. However, regardless of the advantages PSS offer, e.g. in terms of competitive position, customer loyalty and customized product solutions, a large number of SMEs remain operating as traditional producers who sell the physical product as the main sales object and only in some cases offer individual services like repair and spare parts delivery upon request. One of the reasons for remaining with the traditional business model is the lack of time and the lack of technical and methodological resources which are required to overcome the barriers that are linked with the servitization [6,7]. Existing process models and business model patterns that can be used to develop business models for PSS provide a good framework for the development of the integrated services and offer methods for implementation. Yet, the existing process models do not provide a detailed support, especially for SMEs, as they are often designed to be industry-neutral and application neutral, thus rather abstract [8]. Therefore, this paper examines the research question "How does a methodical approach for detailed planning of service delivery processes to offer a PSS has to be designed so that SMEs can independently manage the business model transformation to offer a product solution consisting of the product and integrated services?". For this purpose, chapter two first deals with a general overview of business model development and implementation, existing models, phases and methods and a summarized evaluation with regard to suitability for SME support. The actual methodical approach for detailed planning of service processes for manufacturing companies in business model transformation and its considered development procedure are presented in chapter three. Chapter four then summarizes the methodical approach, considering the research question, and provides an outlook on how it can be expanded and used further.

2. Business model development and implementation – initial situation

A business model generally represents the framework for all entrepreneurial business processes on an abstract level and serves as a management tool with the help of which the business activities can be described, analyzed and designed holistically [9,10]. It describes in an abstract way the architecture of value creation by mapping how a company creates, provides and protects value [11,12]. This can be seen on the middle level of the pyramid, which represents the levels of business model differentiation according to Osterwalder and Pigneur (see Figure 1). The strategy at the planning level determines the direction and scope for business models. The company-specific implementation includes the design of the business processes, which is executed on the implementation level [13,14].

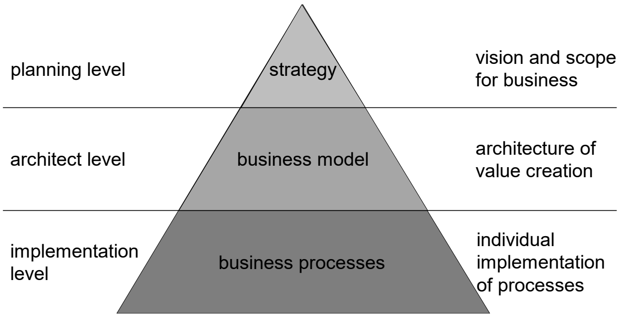


Figure 1: Business model differentiation according to levels and characteristics based on [14]

In order to achieve the strategic goals and taking into account the strategy as orientation as well as the scope for action, the business model development represents a core element for a successful transformation. More and more manufacturing companies are pursuing the goal of developing new business models, thus changing into PSS providers. This transformation requires the implementation of change management processes and innovation processes, which are initiated by means of the methods and tools specified in the process models.

In the literature and in practice, a variety of process models exists to support and guide companies through the development and innovation process. Gassmann et al. for example composed a database and procedure called the *St. Gallen business model navigator* which includes *55 innovative business concepts* [15] to help companies develop new business models. Osterwalder and Pigneur also published a widely known guideline [15]. According to Lins et al., who analyzed different process models, 25 process models were selected from a literature research according to criteria such as, among others, providing real process models and a connection to digitization and industry 4.0 [8]. Within these selected models six general phases (*preparation*, *idea generation*, *design*, *evaluation and selection*, *implementation* and *sustainment*) for business model development were identified [8], which are fully or at least partially addressed by the various models. The start of the transformation process begins in the preparation phase which supports the companies to become aware and prepare for the transformation. This is followed by the creative process for new business models, their design and evaluation. When a new business model is selected it gets implemented, operated and continuously monitored and improved. The whole procedure is an iterative process and needs to be regularly run through by the companies in order to remain competitive. Most of the process models for business model development suggest practical methods and tools to support a successful realization of each process step. Thus, the authors provide tools like the business model canvas from [16] or supporting guiding questions, business model patterns (e.g. [17,15]) or software programs as possible methods [8]. Guiding questions can be used as reference during the transformation process while the business model patterns serve as input for the creative and development phase giving examples on actual business models or samples with abstracted models of different business models of different branches. Software-based tools (e.g. the business model configurator by [17] or the tool for the business model engineering by [4]) support automated the execution of the steps and its methods so that the user can manage the implementation of each of the phases more quickly and easily based on the stored databases and software processes.

The evaluation of the analysis of process models for business model development by [8] shows that the detailed steps of the process models and the linked methods help the companies to understand the different and necessary steps well and to apply them more easily and correctly. It also underlines that business model patterns and best practices are a good orientation for the companies which want to renew their business model or want to develop a new one. But one occurring barrier especially for SME is the final and detailed planning of the processes for the new business model and its implementation because the existing process models are on an abstract level and industry-neutral so that the companies are not able to perform the transfer and the individualization on their own. As long as there is a lack of an individual development methodology with consideration of the existing internal processes and the strategic orientation of the respective company, the process models and its elements that are generally known in science have to be extended accordingly [18]. This can also be applied to the software-based tools, which are a good support but, like the underlying approaches, are on too abstract level to answer the research question. In addition, the approach must be easily and independently implementable by companies and contain solution-patterns that are understandable and can be individualized and transferred to the processes of the companies. In order to provide full support to companies, the solution-patterns and requirements for the chosen scope must be complete [8]. Therefore, in the following a methodical approach for detailed planning of service processes for manufacturing companies is presented, that supports the specific individual realization of service delivery processes for new business models.

3. Approach for detailed planning of service delivery processes in PSS

3.1 Research design and boundary conditions

The approach for detailed planning of services should support companies by planning their service delivery processes in detail and get necessary requirements for the implementation. Therefore, the approach provides a service catalog, which represents the core element. The service catalog includes solution-patterns in form of process modules with stored process models. These solution-patterns are created by applying the theory of inventive problem solving (TIPS also known as TRIZ) according to [19] as the considered research design and are then listed in the catalog. After the catalog got created it can be used by the companies to plan their new service delivery processes in detail. The procedure for the detailed planning by companies is methodically supported by applying the TIPS method with the new created solution-patterns. TIPS originated in Russia and describes a method to solve (technological) problems and innovation tasks. It follows the three hypotheses that abstract problems and solutions repeat independent of their field of knowledge, that similar patterns are the basis for many technical developments and that reasons for innovation are often made from external influences [19]. The theory includes two levels, the concrete and the abstract level. First a concrete problem is abstracted before an abstract solution for the abstracted problem is sought. In a third step the abstracted solution gets transferred to the concrete level and a concrete solution for the origin problem is developed [20]. The abstract description of the problem is carried out by 39 predefined parameters while the abstract solution is generated from 40 defined principles. The TIPS principles and parameters were derived by analyzing several thousand patents and were then summarized in patterns which are represented by the principles and parameters. This means that the inventor of TIPS first sought concrete solutions (patents) for abstract problems, and then abstracted them. By collecting the many solutions and bundling them into patterns, it was then possible by following the known TIPS method to first generate abstract solutions for abstract problems and specify them in the final step.

For the present research question the single parameters and principles are not usable, because they are based on technical contradictions and are not suitable for business model development [20]. But in accordance with the systematic of Lehner, who used this procedure for the solution-pattern-based development of business models for frugal innovation [20], the hypotheses and procedure to solve a problem almost automated with standard solution-patterns that are finally individualized are of huge interest and suitable for the approach of detailed planning. Therefore, the steps of creating the principles and parameters from concrete solutions according to the TIPS inventor will be first used as research design for creating new solution-patterns. The solution-patterns of services for the manufacturing companies consist of derived standardized and generally applicable process models which are combined and linked to modules and these modules to services. The detailed services are then listed in the solution-pattern-based service catalog (see Figure 2). And secondly the actual TIPS method is used as a process model for the companies to plan their service delivery processes in detail by applying the newly created solution-patterns (see chapter 3.3).

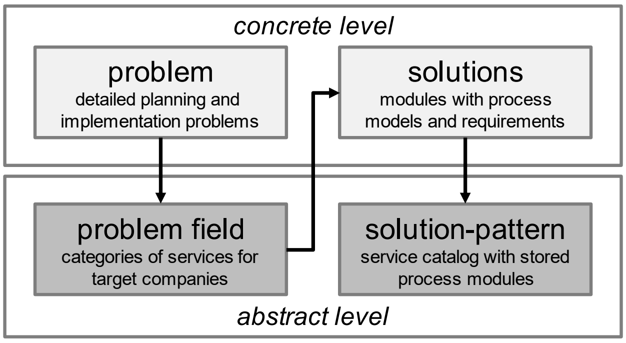


Figure 2: Research design for creating solution-patterns for service delivery processes in adaption to TIPS

To meet the requirements of specific and detailed process modules and process models, that can be used and individualized by companies especially by SME on their own, the methodology considers the following boundary conditions. By the fact that specific process models, which are designed to be realistic and detailed, are necessary as solution-patterns, which can be easily transferred to the own company, the approach defines a clear scope that is considered. In this case the scope is set for companies from the machines and plant engineering. Although, the general procedure of the approach to derive specific process models is applicable for all branches and fields where service processes need to be developed, it is so first implemented for SME from the machines and plant engineering. This scope is set to have a clear area for analyzing which must be completely covered with detailed solution patterns. This focus is also laid of manufacturing companies because as mentioned before their need is to develop integrated services to act as a provider of PSS to stay competitive and in this branch, machines are often still seen as the actual technical products. The still predominantly classical and conservative manufacturing SME particularly need to get support in developing new business models for offering PSS to consolidate or even expand their competitive position in a globalized and digitalized market [21].

3.2 Solution-pattern-based service catalog to offer PSS

The creation of the service catalog following the presented research design considers the two phases identification and derivation of a listing of suitable services (steps “problem” and “problem field” in Figure 3) and creation of solution-patterns as process modules consisting of process models for the detailed planning (steps “solutions” and solution-pattern” in Figure 3).

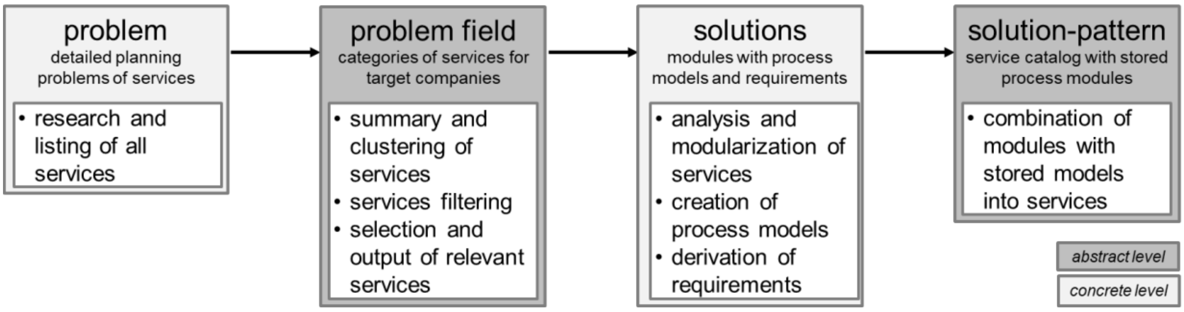


Figure 3: Process steps for creating a solution-pattern-based service catalog

Starting from the concrete problem that companies are not able to develop and implement the necessary service delivery processes for new business models to offer PSS on their own, suitable services are researched and listed. The systematic literature research was conducted in the databases “Web of Science” and “Scopus”. It was carried out with terms which contain combinations of relevant search strings from the area product-service-system, such as *product service system, servitization, industrial product service system, sales service* or *smart services*, and from the area production, such as *machinery, manufacturing* or *production*. Also, the database of “Springer” was used with german search phrases like *after sales service, product-service-system* or *services* in combination with the term i*ndustry*. After removing duplicates and reviewing the titles and the abstracts, 27 works are listed and present the basis for the creation of the listing of suitable services. The researched services are reviewed and then summarized and clustered. It is checked if services have common categories or if they are the same or synonymously used. In a next step the summarized list of services that clusters the services in different categories is filtered. To meet the boundary conditions, it was filtered by the following underlying criteria:

- Is the viewed service a real performance (process) in the sense of a value proposition?
- Is it a performance for which a customer might be willing to pay?
- Is the service standardizable and transferable or only individually for a single company?
- Does the service show a relation to production and manufacturing as well as a relevance for manufacturing companies?

All services were evaluated according to these criteria. Finally, an overview in the form of a listing with the different selected categories and their subordinated services is derived and visualized. Categories corresponding to the criteria are, among others, monitoring, maintenance, operations, spare part management, installation, remote services, and trainings.

In the following the categorized and listed services get analyzed and broken down into individual process modules on a second level and these modules then into process models on a third level with concrete solutions (see Figure 4).

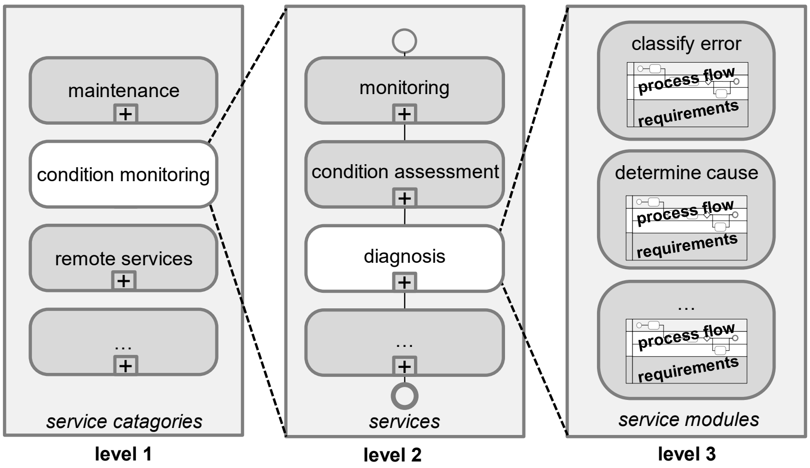


Figure 4: Structure of the service catalog with examples

Each service is reviewed in which modules it can be divided by using standardized descriptions of the services. For this purpose, norms, standards and technical literature (e.g., German Norm DIN, European Norm EN, VDI or VDMA norms) that describe the service or parts of the service processes in a standardized way are used as references. For example the service category *maintenance* can be divided, among others, into the services servicing, inspection, repair and improvement which can be divided into other modules on a more specific level on the basis of the DIN norm 31051 [22]. The subcategory servicing includes the process modules cleaning, readjustment, refilling, replacing and lubrication [22] (exemplary see Figure 5). In addition, more subdivisions according to the planning ahead and the used database are derived for these services and are compatible with the other modules. Following this general procedure, process modules for all services are derived. On a third level there are explicit process models which are assigned to the process modules. As with the subdivision of the listed services, the process models are created by further breaking down the process modules by using standardized descriptions. Each process module can be created by using one or combining more than one process model.

The process models are the core element for the detailed planning. They provide all information about the processes step by step. The sequences are modeled and visualized which is done with the help of the method business service blueprint modeling (BSBM). The BSBM represents an approach for modelling service processes with customer interaction. For this purpose, it combines the standardized and well known notation of the business process modeling notation (BPMN) with the structure and layout of the method service blueprint (SB) [23]. Based on BSBM an uniform and customized template for describing and visualizing the process models is created. Accordingly, all process models are created with a starting point, defined process steps with specified order and necessary branching points, level of customer interaction and end point. Adapted from the SB and in consideration of services for offering PSS the possible forms of customer interaction vary from customer activity over provider activities with customer contact to provider activities in the background without customer contact. The process steps are arranged in the respective lanes by their level of interaction and are linked with arrows in the correct order. Additionally, special requirements for each process step are defined to support the implementation of new service delivery processes for new business models in the companies. The focus of the requirements is on the one hand on the competencies required to perform the process step and on the other hand on the necessary (technical) data and their properties for steps of data-based services. The requirements are noted for each step and are connected to it.

For the required competencies for the technicians who need to execute the process steps a literature research is carried out. As basis for the overview of possible required competencies the CDIO Syllabus v2.0 by [24] is used that defines in detail skills and goals for the engineering education by comparing the first version of the Syllabus with the UNESCO four pillars of learning and with national accreditation and evaluation standards of several nations [24]. These results are then aligned with the other approaches to competence management. The aligned competencies are evaluated against the last criteria that is defined for the evaluation of the services. It is checked if they show a relation to technical processes, production, manufacturing and (technical) data processing. Competencies for the common interaction with customers like e.g., kindness and language, are highly important for a service technician but they are taken for granted and are not part of the present approach. The other matching competencies are summarized in a checklist that facilitates the assignment of competencies to the process steps. Also, a checklist is given for the second group of requirements. Data as well as its correct capture and use are gaining in relevance for all services especially for smart services [25]. Therefore, the checklist mentions possible and exemplary data requirements. Among others, it should propose the type, origin and intake, the quantity and classification, as well as the evaluation of the data as points to be considered for each process step. The checklists for both groups of requirements serve as a help that for all steps the requirements are standardized assigned and that none are forgotten. Finally, a process model defines in detail the necessary process steps, their level of customer interaction and special requirements.

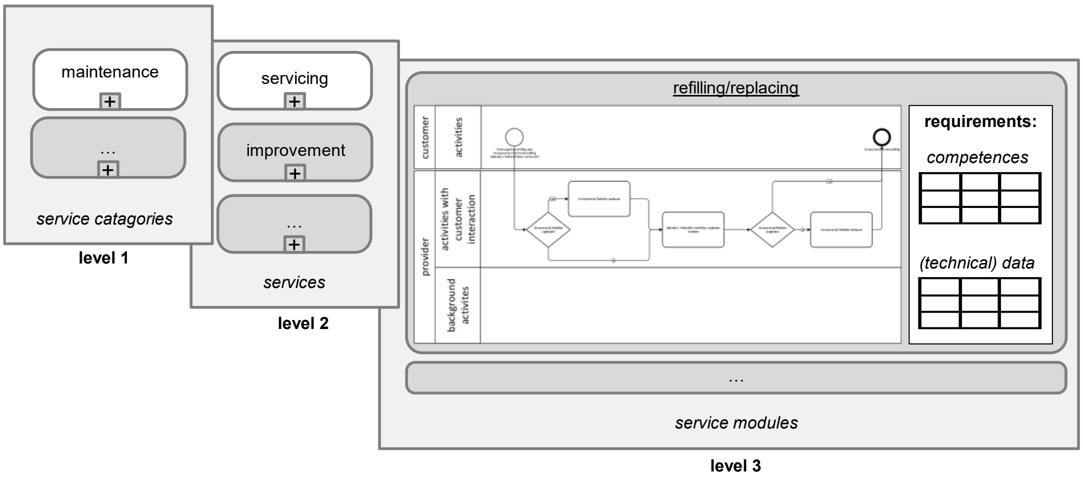


Figure 5: Extract from the modularization of the example "maintenance"

The solution-patterns of services for the manufacturing companies are created by combining the derived process models to modules and these modules to services. With this procedure the detailed and subdivided process models can get linked to one or more modules and services, because many of the generalist process models are part of different service delivery processes. The services and modules are listed in the service catalog as solution-patterns. By the fact that the process models, which contain all information of the required process steps and requirements, get combined to services and are linked to these, there are all information noted and visualized for each service. In this way the solution-pattern-based service catalog for the detailed planning of services to offer PSS includes the model of the ideal-typical process flow composed of the different process models and the requirements related to necessary competencies and data. It represents the input for the actual detailed planning of the service delivery processes by the companies. The catalog can be extended for another scope or other branches by following the steps of the methodology.

3.3 Procedure for implementing service delivery processes by companies

Manufacturing companies which want to plan a new business model in detail and implement it, often get through the business model development phases until they have to develop their individual processes und structures in the companies. That means they create and evaluate ideas for new business models within the scope of the intern strategy and select one or more models for further processing. At that point in many cases they miss time, technical and methodological resources to transfer the ideas into real implementation projects and remain with their traditional business model. Therefore, the second part of the methodical approach is designed to support companies successfully implementing integrated service delivery processes to transform to a provider of PSS using the solution-pattern-based service catalog. The approach foresees to apply the TIPS method with the newly created solution-patterns to get the individual processes for the companies for the chosen service as individual solution (see Figure 6).

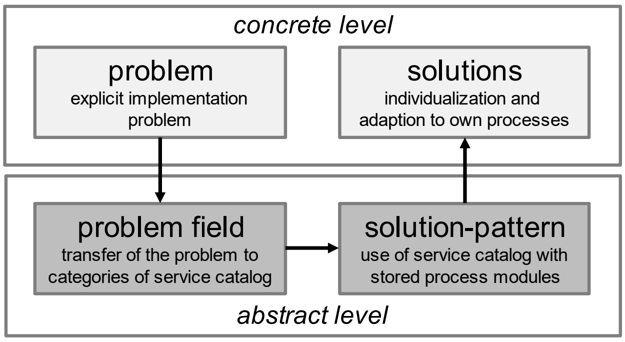


Figure 6: TIPS-based procedure for applying solution-patterns for the detailed planning by companies

Companies which already got through the creative phase of the business model development or already know their desired business model can check out the service catalog for their target service category and target services. By choosing the target service they get the deposited information including the ideal-typical process flow and the requirements related to necessary competencies and data as results which can be used for the next actions by the companies. In a next step they have to individualize and adapt the processes according to the own and individual conditions and situation. Simultaneously a matching and an analysis of the required competencies and data with those available in the company are carried out. At that point the companies do have all information about the required processes with individual process steps that can get implemented and it can be decided whether all requirements are met. If a deviation occurs during the matching and the analysis of the actual and the target situation the approach with its detailed information serves as a decision support e.g., whether to build up missing skills through training or buy them in, whether to outsource the entire process or revise the idea of the selected business model again. This approach enables companies to independently plan in detail and develop the existing or generated ideas for a new business model and thus successfully move from the idea of a business model to its real implementation.

4. Summary and outlook

The transformation and the development of new or adapted business models is a current topic in practice and research supported in particular by the ongoing digitization and industry 4.0. Manufacturing companies are increasingly trying to transform their business model in terms of offering PSS by combining their physical product with integrated services. The business model in general serves at an abstract level as a management tool for visualizing, analyzing and developing all business processes required for value creation. Therefore, there is a variety of process models that aim at the development of business models and offer methods and tools for the transformation process. Although many methods and tools are mentioned there is a lack of support for the competence-based detailed planning and individual implementation of the new business models developed in the creative phases. This concerns especially SMEs, some of which do not have the necessary resources to perform that transfer on their own. This gap is addressed with the present methodical approach which is based on the solution-pattern-structure of the theory of inventive problem solving. The methodology foresees two procedures. The first describes the creating of the service catalog with process modules and linked process models. The catalog provides information about the ideal-typical process flow and requirements for the process steps because each service delivery process is formed by a combination of different process modules, which in turn are composed of the derived elementary process models. The second procedure addresses the application of the service catalog by the companies to plan their business models in detail and implement the corresponding service delivery processes on their own. The approach provides relevant services with ideal-typical, detailed process models and required competencies, so that companies from the mechanical and plant engineering sector can directly test and independently implement the delivery processes of the new PSS. The procedure for creating the solution-pattern-based service catalog is generally valid and transparent, so that it can also be transferred and elaborated in further work to areas and service processes outside the set boundary conditions. Also, the requirements elaborated here, which are based on the boundary conditions, can be extended according to the needs of the target group in terms of the present procedure. In this way, the service catalog can be extended and can be used in fields and branches other than its original narrow focus. In future work a proof of concept must be further carried out with manufacturing companies for the application of the methodical approach for detailed planning of services to offer PSS with the specific focus set by the boundary conditions. For user-friendly use, the contents of the service catalog, such as categories, modules, process models and requirements as well as dependencies, will be transferred into a software-based tool.

Acknowledgements

Parts of this article have been elaborated during the research and development project ABILITY (reference number: 02L17B024), which is financed with funding provided by the Federal Ministry of Education and Research and the European Social Fund under the "Future of work" program.

References

[1] Herrmann, K., Werkle, M., Prinz, C., Lins, D., Arnold, D., Köhler, C., Mahl, T., 2021. Hybride Wertschöpfungspotenziale in kleinen und mittelgroßen Unternehmen. Bundesministerium für Arbeit und Soziales (BMAS). https://www.esf.de/SharedDocs/Meldungen_NL/Newsletter/2021/nl_ability.html. Accessed 25 January 2022.

[2] Coreynen, W., Matthyssens, P., Van Bockhaven, W., 2017. Boosting servitization through digitization: Pathways and dynamic resource configurations for manufacturers. Industrial Marketing Management 60, 42–53.

[3] Kagermann, H., Wahlster, W., Helbig, J. (Eds.), 2013. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0. Deutschlands Zukunft als Produktionsstandort sichern, Frankfurt/Main.

[4] Boßlau, M., 2014. Business Model Engineering: Gestaltung und Analyse dynamischer Geschäftsmodelle für industrielle Produkt-Service Systeme. Dissertation, Bochum.

[5] Lins, D., Arnold, D., Prinz, C., Kuhlenkötter, B., 2019. Befähigungssystem für die Transformation zu hybrider Wertschöpfung. ZWF 114 (12), 851–854.

[6] Lins, D., Arnold, D., Mahl, T., Köhler, C., Kuhlenkötter, B., Prinz, C., 2021. Phasenmodell zur Überwindung von Implementierungsbarrieren bei der Entwicklung hybrider Geschäftsmodelle, in: Gesellschaft für Arbeitswissenschaft e.V. (Ed.), Arbeit HUMAINE gestalten. Bericht zum 67. Arbeitswissenschaftlichen Kongress. GfA-Press, Dortmund.

[7] Wiesner, S., Nilsson, S., Thoben, K.-D., 2017. Integrating requirements engineering for different domains in system development – lessons learnt from industrial SME cases. The 9th CIRP IPSS Conference: Circular Perspectives on Product/Service-Systems 64, 351–356.

[8] Lins, D., Arnold, D., Köhler, C., Mahl, T., Prinz, C., Kuhlenkötter, B., 2021. Analysis of process models for the business model development considering special SME requirements for offering PSS, in: Proceedings of the 2nd Conference on Production Systems and Logistics (CPSL 2021). 2nd Conference on Production Systems and Logistics. publish-Ing., pp. 108–117.

[9] Köster, O., 2014. Systematik zur Entwicklung von Geschäftsmodellen in der Produktentstehung. Verlagshaus Monsenstein und Vannerdat OHG, Münster.
[10] Zollenkop, M., 2006. Geschäftsmodellinnovation: Initiierung eines systematischen Innovationsmanagements für Geschäftsmodelle auf Basis lebenszyklusorientierter Frühaufklärung, 1st ed. Deutscher Universitäts-Verlag GWV Fachverlage GmbH, Wiesbaden.
[11] Schallmo, D., 2013. Geschäftsmodell-Innovation: Grundlagen, bestehende Ansätze, methodisches Vorgehen und B2B-Geschäftsmodelle. Springer Gabler, Wiesbaden, 335 Seiten.
[12] Wirtz, B.W., 2013. Business Model Management: Design - Instrumente - Erfolgsfaktoren von Geschäftsmodellen, 3., aktuelle und überarbeitete Auflage ed. Springer Gabler, Wiesbaden, 400 pp.
[13] Casadesus-Masanell, R., Ricart, J.E., 2010. From Strategy to Business Models and onto Tactics. Long Range Planning 43, 195–215.
[14] Osterwalder, A., Pigneur, Y., 2002. An e-Business Model Ontology for Modeling e-Business, in: Proceedings of 15th Bled eConference e-Reality – Constructing the e-Economy, Bled, Slovenia. June 17-19 2002, pp. 1–12.
[15] Gassmann, O., Frankenberger, K., Csik, M., 2013. Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Hanser, München, 320 pp.
[16] Osterwalder, A., Pigneur, Y., 2011. Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag, Frankfurt, New York, 282 pp.
[17] Echterhoff, B., Koldewey, C., Gausemeier, J., 2017. Pattern based business model development - identification, structuring and application of business model patterns. The ISPIM Innovation Forum.
[18] Peruzzini, M., Marilungo, E., Germani, M., 2014. Functional and Ecosystem Requirements to Design Sustainable Product-Service, 768–777.
[19] Echterhoff, N., 2014. Systematik zur Planung von Cross-Industry-Innovationen. Dissertation, Paderborn, 210 pp.
[20] Lehner, A.-C., 2016. Systematik zur lösungsmusterbasierten Entwicklung von Frugal Innovations. Verlagshaus Monsenstein und Vannerdat OHG, Münster.
[21] Luedeke, T.F., Köhler, C., Conrad, J., Grashiller, M., Ruf, T., Sailer, A., Vielhaber, M., 2018. CPM/PDD as an integrated product and process model for a design-thinking based, agile product development process, in: , Proceedings of the International Design Conference - Design 2018. Dubrovnik/Croatia, May 21 - 24, pp. 2063–2074.
[22] Deutsches Institut für Normung e.V., 2019. Grundlagen der Instandhaltung. Beuth Verlag GmbH, Berlin 01.040.03; 03.080.10. Accessed 12 October 2021, 13 pp.
[23] Meis, J., Menschner, P., Leimeister, J.M., 2010. Modellierung von Dienstleistungen mittels Business Service Blueprinting Modeling, in: Thomas, O., Nüttgens, M. (Eds.), Dienstleistungsmodellierung 2010. Physica-Verlag HD, Heidelberg, pp. 39–64.
[24] Crawley, E.F., Malmqvist, J., Lucas, W.A., Brodeur, D.R., 2011. The CDIO Syllabus v2.0. An Updated Statement of Goals for Engineering Education, in: Proceedings of the 7th International CDIO Conference. Technical University of Denmark, Copenhagen, Denmark. June 20 – 23, 2011.
[25] Koldewey, C., Meyer, M., Stockbrügger, P., Dumitrescu, R., Gausemeier, J., 2020. Framework and Functionality Patterns for Smart Service Innovation. Procedia CIRP 91, 851–857.

Biography

Dominik Lins (*1991) has been working as a research assistant at the Chair of Production Systems (LPS) at the Ruhr-University Bochum since 2016 in the field of production management. He earned a bachelor’s and master’s degree in mechanical engineering at the Ruhr-University Bochum. His primary research topics are the digitalization of production systems and product service systems.

Until 2009 **Bernd Kuhlenkötter** (*1971) was responsible for product management and technology at ABB Robotics Germany. In 2009 Bernd Kuhlenkötter took over the Professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Since 2015 he holds the professorship of the Chair of Production Systems (LPS) at the Ruhr-University Bochum and is managing director of the Institute for the Engineering of Smart-Product Service Systems.

3rd Conference on Production Systems and Logistics

Process Data Validation For Manual Assembly Systems Used For Highly Variable Products

Martin Sudhoff[1,]*, Jonas Viehöver[1], Michael Herzog[1], Bernd Kuhlenkötter[1]
[1]Ruhr-Universität Bochum, Chair of Production Systems, Universitätsstr. 150, 44801 Bochum, Germany

Abstract

The production of customized goods is becoming more and more important for industrial companies. The large number of variants resulting from this, up to batch size 1 production, requires a high degree of flexibility. To meet these requirements, manual production processes are frequently still used. This is especially applicable to the area of assembly. Data acquisition is a significant task in manual assembly due to volatile secondary activities and alternative handling operations. The process times to be recorded are also influenced both consciously and unconsciously by the employees. This paper describes an approach for the validation and interpretation of production data of manual assembly systems. Therefore, process data are analysed based on the use case of terminal strip assembly in the learning factory of the Chair of Production Systems at the Ruhr-University Bochum is presented. Here, the validation of the product data from 2021 is carried out by checking the data for normal distribution. This is followed by an analysis of the data with regard to the effects of spikes. Furthermore, the influences of a low data basis, different degrees of standardization and learning effects in the course of production are analysed. Finally, a discussion on the findings and further procedures will take place.

Keywords

Process time acquisition; Production planning; Manual assembly; Factors of influence; Validation

1. Introduction

Industrial production is increasingly subject to the change from a supplier's to a buyer's market. Product life cycles are becoming shorter and the number of product variants is increasing. [1, 2] In addition, labour costs are rising and demographic developments are leading to a shortage of skilled workers [3]. In connection to the increasingly demanded of high flexibility and adaptability, this means a great challenge for manual assembly. In the field of manual assembly, there is still no automated acquisition and real-time-capable evaluation of production data for dynamic production control [4, 5]. One reason for this is that manual operations are often used where tasks change volatilely and a high degree of adaptability is required. In addition to these workflows, which are rather difficult to predict, individual approaches of people also predominate. Often, a lack of standardization in SME production environments is an additional factor here. Finally, data acquisition must not mean any additional work for employees, in order not to increase the proportion of secondary activities. Furthermore, the privacy rights of employees must be given priority.

In this conflict, both the continuous data acquisition and subsequent filtering as well as the interpretation of the KPIs do not mean a proper representation of the production system. In addition to the high process diversity and changing environmental conditions, the diverse human factors, such as motivation, learning

DOI: https://doi.org/10.15488/12165

behaviour and fatigue, pose a particular challenge for analysis and conclusions. Especially the acceptance of production data acquisition systems in production environments dominated by humans requires accurate data and key figures. With regard to the validation and analysis of data, there are approaches in the area of automated machines. The data captured automatically by means of sensors can be examined with regard to trends and distribution functions. Furthermore, approaches such as the DOE are state of the art in the analysis of technical systems. [6] Here, individual parameters of the systems can be varied and examined with regard to their model. This is not possible in the real production process of a manual assembly.

For this reason, in this paper an approach for the validation and interpretation of production data of manual assembly systems is presented. Therefore, the state of the art in the field of data acquisition, processing and evaluation in manual assembly is discussed first. Then, process data are analysed based on the use case of terminal strip assembly. Here, the validation of the product data from 2021 is carried out by checking the data for normal distribution. This is followed by an analysis of the data with regard to the effects of spikes. Furthermore, the influences of a low data basis, different degrees of standardization and learning effects in the course of production are analysed. Finally, a discussion on the findings and further procedures will take place.

2. Process time acquisition in manual assembly systems

Manual assembly systems are often used where tasks change volatilely and a high adaptability to the changing circumstances is required. At the same time, due to this high flexibility, there is a challenge to be able to plan and control this type of work process efficiently with regard to productivity targets. [4, 5] In contrast to automated production, manual assembly systems are subject to a large number of influencing factors which have an impact on the assembly time as an important KPI. As illustrated in Figure 1, in addition to product and process properties as well as the design of the supplier network and the assembly system with its environmental conditions, primarily human factors influence the actual execution of the assembly task.

In the context of production planning, an assembly process is initially derived from the existing p*roduct properties* and the design of the manufacturing system. In relation to the product, its complexity should first be mentioned as an important factor influencing the process time. The authors Samy/ElMaraghy [8] define a significant correlation between increasing product complexity and increased assembly effort. The complexity measure described here depends, among other things, on the number and geometries of the components. Other important factors are the required tolerances and the type of product structure. [9] The *process properties* are directly dependent on the product properties. One of the most important influencing factors is the level of standardization [10, 11]. The manual processes are often found in SMEs. Especially in these companies with only a few employees, the workflows are less standardized and the workers often have a more diversified range of tasks. The lower the degree of standardization in the assembly system, the greater the possibilities for individual operations by the employees. This makes it more difficult to capture operating data accurately. Furthermore, the scope of the process and the process complexity are relevant factors [10, 11]. Within the context of assembly, this complexity results from energetic and informational activities. According to Schlick, assembly involves precise movements with low forces and can therefore be classified as a rather energetic task with an informational proportion. [12] In particular, the technological and content-related process diversity as well as the number of assembly operations are major complexity drivers [13]. In the category of *production system*, the number of workstations and their layout have a major influence [1, 13]. In this context, the differentiation between one-piece flow and batch production is also important [1]. Finally, the required equipment and supplies [13], as well as ergonomic aspects, are also significant.

Subsequently, the *environmental effects* of the assembly system are also relevant for the work being performed. These include first and foremost the fundamentals defined in DIN 6385 "Principles of ergonomics for the design of work systems" with regard to the aspects of temperature, light or even air purity.

[14] Also, the framework conditions related to *scheduling and network* aspects have an influence on the actual activity carried out. The order volume [10] and the order sequence need to be mentioned as well. Furthermore, delivery dates and the delivery reliability of suppliers in particular influence the work of the employees. This also includes the number of suppliers and the quality of the components [13].

In addition to these technical influencing factors, it is primarily the *human factors* that have an impact on the assembly processes actually carried out in manual production systems. In the context of this examination these are classified into the categories of qualification, learning, stress, fatigue, motivation and well-being. [10, 13] In addition to basic qualification [15], learning in production is a particularly important influencing factor [16, 17]. In the production environment, learning usually means a decrease in processing times and material consumption over time. This occurs due to repetitive work operations and the increasing experience of the employee as a result. This relationship was published by Wright as learning curve theory. Assuming unlimited time and a constant learning rate, it can be observed that the average cumulative value per product decreases by the same rate when the number of products is doubled. [16, 12] In line with current knowledge, this correlation has been adjusted so that an average learning curve for a batch can be characterized by processing times which initially decrease steeply and then more slowly as the number of units increases. Based on the non-constant learning rate, a level of saturation finally results. Such learning effects occur especially at the beginning of a batch, which is why they are of particular importance in the production of small batches [18]. Furthermore, the more complex the activity, the steeper the learning rate [19]. In addition to the learning effects, the aspects of stress and fatigue [10] as well as motivation and well-being also have an influence on the assembly time [20]. These aspects are based on the four levels of Maslow's pyramid of needs. This was further developed by Landau according to production-specific issues [20]. Thus, the activity should first be theoretically feasible. Based on this, Landau describes the tolerability that can be achieved by designing occupational safety according to the state of the art. Finally, an activity that can be performed on a permanent basis is expected to be reasonable. This third stage involves a human-oriented work design as well as a fulfilment of the employees' expectations. Finally, the goal of the fourth stage is to achieve a high level of satisfaction by ensuring the development of personality as well as social acceptance. [20, 21]

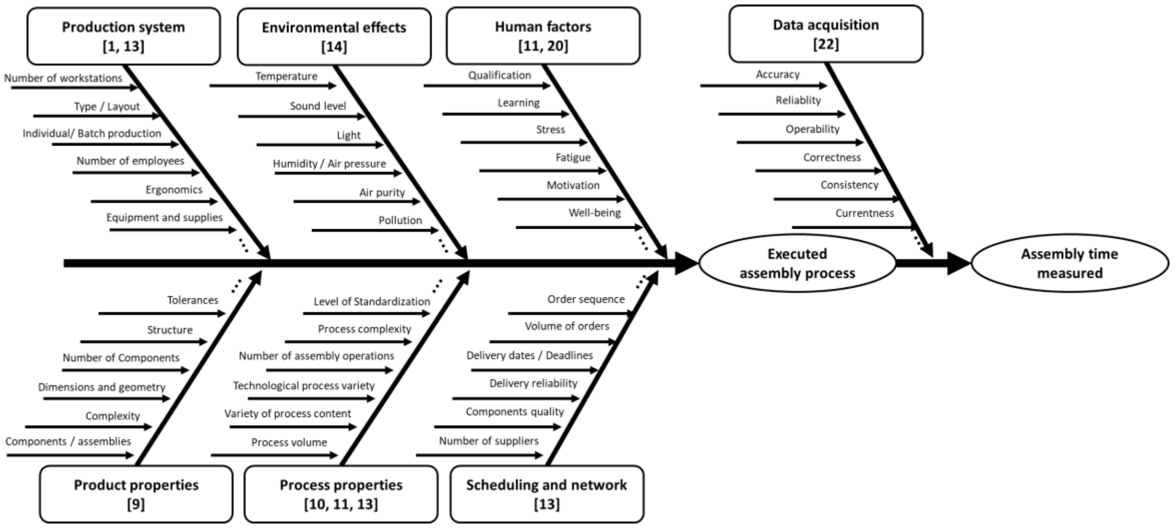


Figure 1: Factors of influence in manual assembly systems

Considering this complex relationship, it is not surprising that it can be seen in practice how the actual process deviates in principle from the specifications of the assembly planning. Even if the same target of assembly of two or more components is achieved, it is still an individual variation of the specified work process. [7] Therefore, real data from the assembly system are essential for efficient production control in terms of production data acquisition. From the diverse constraints in manual assembly, a high level of complexity can be derived for the design and implementation of a production data acquisition system (PDA).

This must consider the volatile activities as well as the fluctuating environmental conditions and at the same time must not disrupt the workflow or lead to increased secondary workloads. Still, it must be observed that no supervision of the employees takes place and at the same time an objective data basis is ensured. [4] Here, the principles of the data quality requirements defined by Fox, such as accuracy and reliability, must be considered [22]. Also due to these aspects, the most commonly used methods for capturing process times are the direct conversation with employees and the use of forms and reports by means of simple software tools [23]. In addition to self-recordings, multi-moment measurements and, time studies according to REFA continue to be the most common time determination methods in the industrial environment [24]. These are usually associated with a large initial effort as well as expenses in case of changes. The acquired data must then be processed. First, outliers and implausible values must be filtered out, and then a validity check must be performed. After Dekena, the box-plot method is suitable for filtering data in a production context [25].

In order to examine the validity of random variables with a continuous probability function, different tests are suitable. Common methods are the $\chi 2$-test, the KS test or the Cramér-von-Mises test. [26] The procedure includes first the choice of a suitable model. Furthermore, the parameters of the model are estimated on the basis of observations and graphical methods [26]. Thus, certain probability distributions can be inferred. In this case, the model of normally distributed random variables is of special importance [27]. The normal distribution is usually seen in populations and can be used to describe the random dispersion of measured values. Besides determinations in natural sciences and medicine, this unimodal distribution is often strongly asymmetric, especially in the context of mechanical engineering. The normal distribution is usually seen in populations and can be used to describe the random dispersion of measured values. Besides determinations in natural science, this unimodal distribution is often strongly asymmetric, especially in the context of mechanical engineering. In the case for positive, right skewed distributed data, the natural lognormal distribution is applied. Here, the higher frequencies are located on the left side. [28, 30] In manual assembly systems, many small random influences overlay each other multiplicatively. This leads to the assumption that, from a certain amount of data, a normal distribution can be observed in the process times. Furthermore, since each assembly task is subject to a strictly physiological limit, a right-skewed distribution is to be expected. This corresponds to a logarithmic distribution. Thus, a variable to be examined is log-normally distributed exactly when its logarithm is normally distributed [29]. For verifying the log-normal distribution, the data values are first logarithmised and then the normal distribution is tested [30].

In this context, the Kolmogoroff-Smirnoff test is used to examine the normal distribution. With this goodness-of-fit test, the empirical distribution function is compared with the theoretical normal distribution. The advantage of this test is the lower effort and a good result even in case of a small number of data values compared to the $\chi 2$-test [31]. This test can be used to check whether a random variable follows a previously assumed probability distribution [32]. In order to verify for normal distribution, the maximum perpendicular distance of the cumulative values is compared with a critical value. The calculation of the critical value depends on the significance level and number of samples [31]. In the literature, the significance level is often set between 0.01 and 0.1 [33]. It defines the range of rejection. When the sample is within this range, the null hypothesis of normal distribution is rejected. Accordingly, the smaller α is chosen, the greater the probability for the result of the investigation to define the data set as normally distributed [34].

3. Examination by means of the case study terminal strip assembly

The approach for validating and interpreting production data of manual assembly systems is explained below using the use case of terminal strip assembly of the Chair of Production Systems. First of all, it must be determined whether the use of the production data acquisition system has resulted in valid production data. This is the basis for using the data for the calculation of KPIs and production control. Furthermore, the data are examined regarding to the influencing factors. In addition to the expected log-normal distribution of the process values, this includes, the analysis of learning curves as well as malfunctions and measurement errors.

3.1 Experimental setup and test procedure

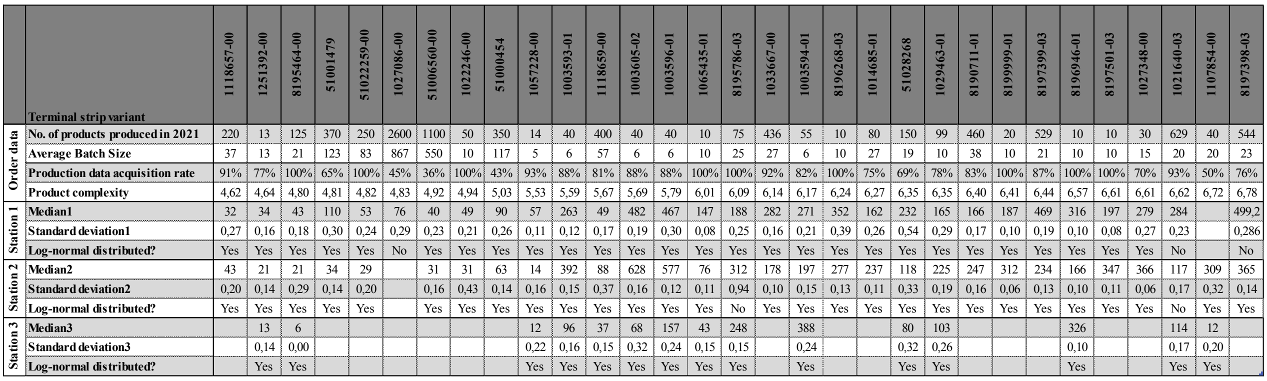
The terminal strip assembly can be assigned to the area of small parts assembly with a high diversity of variants. It is exemplary for other manual assembly processes in mechanical engineering. The assembly system is operated in cooperation between the Chair of Production Systems and the company Phoenix Contact GmbH & Co KG and represents an industrial environment with real orders. [35] The first step in the production of terminal strips is to cut the rails to length. These are then transferred to workstation 1, where the terminals are mounted on the rails. This is followed by the labelling at workstation 2, where small labels are applied. After that, the assembly of circuit bridges takes place at workstation 3. Subsequently, the desired functionality of the terminal strip is ensured by quality tests at workstation 4. Workstation 5 is used for pre-wiring the terminal strips. The final workstation 6 serves to package. [35] The assembly is carried out by two experienced employees who are supported by an assistant during peak loads. The process times of the assistant are not considered. The product portfolio of the terminal strip production comprises 70 variants, each consisting of a unique composition in terms of the number and variation of terminals, labels, circuit bridges and other components. In 2021, 31 of these variants were produced between 10 and 1173 times.

In the scope of this examination, the production period from 07.01.2021 to 03.12.2021 is considered. A total of 8799 terminal strips were produced, of which 5740 were recorded by the PDA. This results in an acquisition rate of 65.23%. The reasons for the non-recorded values of about one third of the products consist in the non-consideration of the process times of the assistant as well as in technical aspects during the introduction and smaller revision steps of the PDA. Ultimately, the familiarization of the employees with the new system, especially at the beginning, also meant for a lower acquisition rate. The processing times were recorded for each assembly part process using a tablet-based app. The processing times for every assembly station are recorded for each order and product. In addition to these productive times (operating state *production*), the pause times and non-productive times, such as setup, rework or malfunctions, were also recorded. In the period under review, 85% of the data relates to the operating state *production*, 10% to *pause* times and 3% to *setup*. The data volumes of the other operating states comprise less than 1%. The subsequent validation of the production data focuses on the *production* times of the terminal assembly, labelling and bridge assembly stations, since this is where the assembly activities take place. In this way, a total of 63% of the productive times of the assembly system are analysed.

With regard to the general conditions under which the investigations were carried out, it can be stated that the influencing factors shown in Figure 1 with regard to the categories of *process properties*, *production system* and *environmental effects* were constant during the period of investigation. The *product properties* are liable to the variance that is defined by the range of parts. In relation to the category *order planning and network*, the factors part quality and number of suppliers are fixed. In the course of real production, changes occur continuously with regard to order-specific aspects such as order sequence, order volume, deadlines and delivery reliability. The changes in these parameters can be considered in the analysis with the help of the PDA and digital order planning. The category of *human factors* is constant in terms of qualification. The parameters learning, stress, fatigue and motivation continue to fluctuate depending on the current boundary conditions and individual constitutions of the employees. These cannot be measured directly. However, conclusions can be drawn from the collected process data, so that indirect statements can be made about this during data analysis. Furthermore, the aspect of well-being cannot be measured with the used PDA system.

3.2 Presentation of results

Consequently, the processing times at the three stations are examined for validity on the basis of the 31 different terminal strip variants. For this purpose, the entire data set is first filtered using the box-plot method. Figure 2 initially shows that the variants were produced in varying numbers and frequencies in 2021. Furthermore, there is an data acquisition rate for each variant. The terminal strip variants are sorted in ascending order with regard to their assembly complexity. The complexity depends primarily on the number of different components, the total number of components and their properties. [36] A more complex variant tends to be more complex to assemble, which is also reflected in the processing times. The average processing times per station is 201 s for terminal strip assembly (labelling: 202 s, bridge assembly: 114 s).



	Terminal strip variant	1118657-00	1251392-00	8195464-00	51001479	51022259-00	1027086-00	51006560-00	1022246-00	51000454	1057228-00	1003593-01	1118659-00	1003605-02	1003596-01	1065435-01	8195786-03	1033667-00	1003594-01	8196268-03	1014685-01	51028268	1029463-01	8190711-01	8199999-01	8197399-03	8196946-01	8197501-03	1027348-00	1021640-03	1107854-00	8197398-03
Order data	No. of products produced in 2021	220	13	125	370	250	2600	1100	50	350	14	40	400	40	40	10	75	436	55	10	80	150	99	460	20	529	10	10	30	629	40	544
	Average Batch Size	37	13	21	123	83	867	550	10	117	5	6	57	6	6	10	25	27	6	10	27	19	10	38	10	21	10	10	15	20	20	23
	Production data acquisition rate	91%	77%	100%	65%	100%	45%	36%	100%	43%	93%	88%	81%	88%	88%	100%	100%	92%	82%	100%	75%	69%	78%	83%	100%	87%	100%	100%	70%	93%	50%	76%
	Product complexity	4,62	4,64	4,80	4,81	4,82	4,83	4,92	4,94	5,03	5,53	5,59	5,67	5,69	5,79	6,01	6,09	6,14	6,17	6,24	6,27	6,35	6,35	6,40	6,41	6,44	6,57	6,61	6,61	6,62	6,72	6,78
Station 1	Median1	32	34	43	110	53	76	40	49	90	57	263	49	482	467	147	188	282	271	352	162	232	165	166	187	469	316	197	279	284		499,2
	Standard deviation1	0,27	0,16	0,18	0,30	0,24	0,29	0,23	0,21	0,26	0,11	0,12	0,17	0,19	0,30	0,08	0,25	0,16	0,21	0,39	0,26	0,54	0,29	0,17	0,10	0,19	0,10	0,08	0,27	0,23		0,286
	Log-normal distributed?	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No		No
Station 2	Median2	43	21	21	34	29		31	31	63	14	392	88	628	577	76	312	178	197	277	237	118	225	247	312	234	166	347	366	117	309	365
	Standard deviation2	0,20	0,14	0,29	0,14	0,20		0,16	0,43	0,14	0,16	0,15	0,37	0,16	0,12	0,11	0,94	0,10	0,15	0,13	0,11	0,33	0,19	0,16	0,06	0,13	0,10	0,11	0,06	0,17	0,32	0,14
	Log-normal distributed?	Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Station 3	Median3		13	6							12	96	37	68	157	43	248		388			80	103				326			114	12	
	Standard deviation3		0,14	0,00							0,22	0,16	0,15	0,32	0,24	0,15	0,15		0,24			0,32	0,26				0,10			0,17	0,20	
	Log-normal distributed?		Yes	Yes							Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes			Yes	Yes				Yes			Yes	Yes	

Figure 2: Evaluation results

As shown in Figure 2, the procedure for validating process data of manual assembly systems described in Chapter 2 results in a normal distribution rate of 90% for the terminal assembly station. Out of the 30 variants considered here, the distributions of the recorded assembly times of 27 variants are log-normally distributed. A similarly high rate of 93% was obtained for the station labelling. Here, sufficient data were also collected for 30 variants to check the normal distribution. Thus, 2 variants are not log-normally distributed. The third assembly station, bridge assembly, in contrast, has significantly less data. Here, all recorded production times are log-normally distributed.

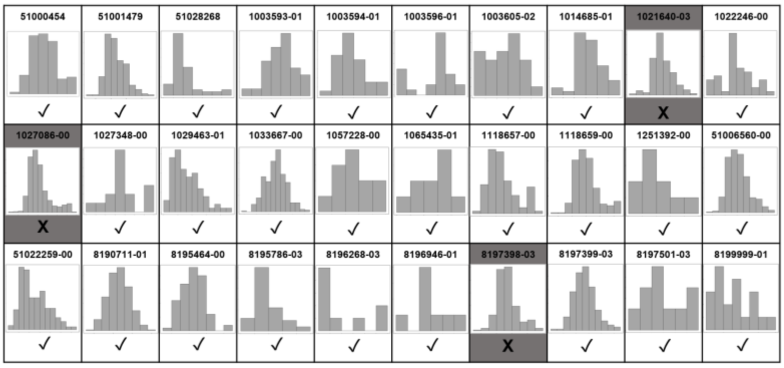


Figure 3: Processing time histograms of terminal assembly (**x** - not log-normal distributed)

3.3 Evaluation and discussion

When taking a closer look at the individual histograms, a bell-shaped or right-skewed, logarithmic distribution can often be seen (see Figure 3). This pattern of the distributions allows the conclusion that it is valid process data. Furthermore, conclusions can be drawn about the production process and the data acquisition method. The frequently observed right-skewed distributions of the process values are shown as an example in Figure 4 using the histogram for the data of variant 8197399-03 of the station labelling. In relation to this variant, 529 products were produced in the period under review with an average batch size of 21 pieces. The acquisition rate is 87% and the median processing time is 234 seconds. The exemplary distribution of the measured values, which is characteristic for the majority of the data, can be distinguished on the basis of three different areas. It should be noted that these areas are not subject to a strict separation, but rather a flowing transition.

The *first area* contains the lowest values up to 200 seconds. Here, a strong increase in frequency can be seen. In connection with the median of 234 seconds, this leads to the conclusion that the majority of the values in this area must be the result of technical measurement errors or incorrect usage of the PDA, since a processing time lower than 200 seconds is not achievable even for a skilled worker. This "physiological limit" can thus be seen in the majority of the histograms. In contrast, realistic values can be assumed in the *second area*. This area contains the majority of the recorded values and always includes the median of the processing time. Nevertheless, there is also a scattering of values between 200 and 260 seconds. This can be interpreted as normal performance fluctuation in manual assembly. In this example, the *third area* of the histogram contains all process values from 260 seconds and includes significantly fewer data values compared to the second area. For the majority of the recorded distributions, a staircase-like decrease in the frequencies with increasing processing time can be seen here. This is probably due to disruptions in the production process and problems with the assembly task. However, data resulting from incorrect operations can also be found here. Another reason for increased process times in this area can be learning effects. In the context of this variant with a high production frequency, this means a rather small influence.

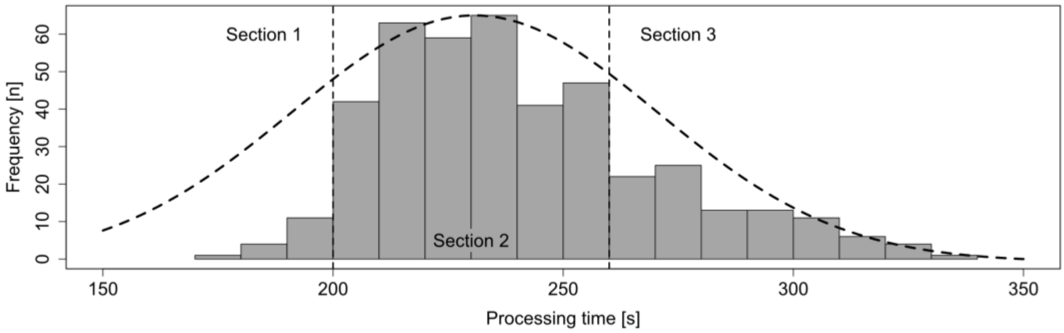


Figure 4: Histogram of the processing times of the terminal strip variant 8197399-03 at the labelling station

In contrast to this example, five data records do not correspond to the log-normal distribution. A closer look at the measured values shows specific reasons, which will be discussed in the following using the four categories *incorrect usage of the PDA, average processing time, lack of standardization* and *low data basis*.

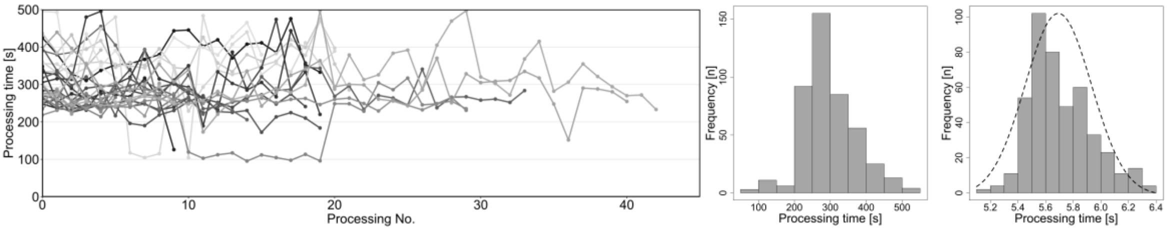


Figure 5: Terminal assembly of variant 1021640-03: a) Production order processing times; b) Histogram; c) Log-normal distribution

First of all, as with the explanation of the measured values from the first and third area, the *incorrect usage of the PDA* should be mentioned as a source of error. This can be seen particularly in the example of variant 1021640-03 (see Figure 5). The main reason for the non-existent log-normal distribution here is the process data of a few orders. This becomes clear when looking at Figure 5. Here, all 31 completed jobs of this terminal strip variant at the terminal assembly station are plotted. As also shown in the histogram, the majority of the measured values are in the range of 200 to 400 seconds. The median is 284 seconds. However, there are also 2 jobs with many physiologically unrealistic values below 200 seconds and a large number of process values above 400 seconds. In total, this means that no log-normal distribution prevails here.

The second category is the *average processing time* of the assembly activity. This is particularly evident in the example of the labelling of the terminal strip variant 102640-03 (see Figure 6). Here, the median of the processing time amounts 117 seconds. Basically, the histogram shows a similar distribution to the example in Figure 5. Nevertheless, the difference here is the significantly shorter processing time, which means that disruptions in the operating sequence and operating errors have a greater influence. With long processing times of several minutes per product, these are less significant than with such short processing times.

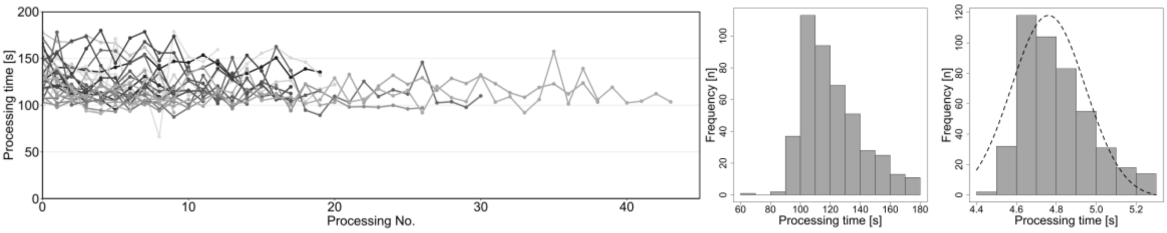


Figure 6: Labelling of 1021640-03: a) Production order processing times; b) Histogram; c) Log-normal distribution

The third category includes the error source of *lack of standardization* (see Chapter 2), which can be seen in particular in the example of the assembly of terminal strip variant 8195786-03. Figure 7 shows that two out of three production orders recorded have a clearly different average processing time. The difference in the processing time corresponds approximately to the duration of the processing time for bridge assembly. In both cases, no data is given for the bridge assembly, which leads to the conclusion that the employees have assembled the bridges already at station 2. This correlation is noticeable because of the few orders with little data, so that the box plot filtering does not apply.

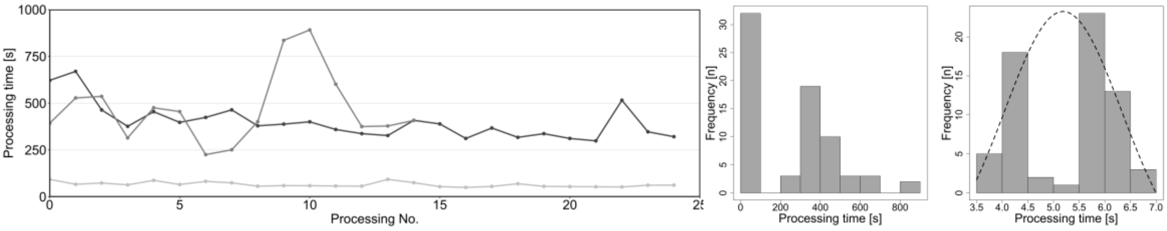


Figure 7: Labelling of 8195786-03: a) Production order processing times; b) Histogram; c) Log-normal distribution

The fourth category includes this context of a *low data basis*. The terminal assembly of variant 1027086 is selected here as the example (see Figure 8). Here there are three different orders with an average batch size of 867. The median processing time is 76 seconds. An examination of the histogram shows that the malfunctions and failures are much more significant here than in the case of more complex variants. Another reason for the larger number of high process values could also be a learning effect, since this variant was only ordered three times.

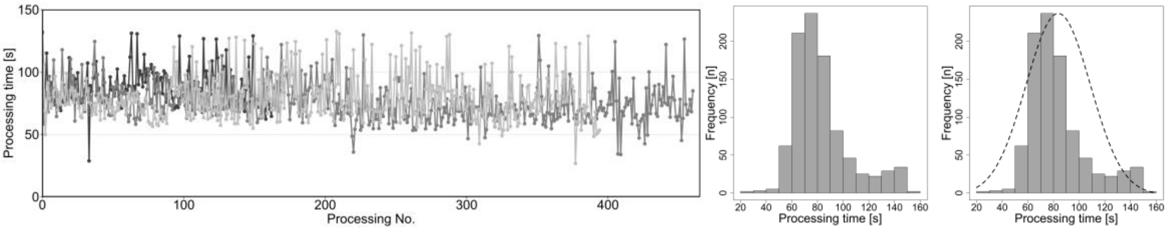


Figure 8: Terminal assembly of variant 1027086: a) Processing times of the individual production orders; b) Histogram; c) Log-normal distribution

For further investigation of possible learning effects, an exponential regression line was calculated for the respective data of the production orders. If the algebraic sign of the gradient is negative for these functions, it can initially be concluded that there has been an average improvement in processing times in relation to the individual products of a batch. Thus, in 75% of the 465 production orders examined, a reduction in the processing time per product can be observed in the course of the work progress. This can be differentiated in relation to the individual stations. This shows that 68% of the orders for terminal assembly have a reduced assembly time in the course of processing (78% for labelling, 86% for bridge assembly). There may be various reasons for this. According to Buck, in addition to the learning effect, changed work processes can be mentioned here, for example. Thus, learning effects cannot be proven here, but only disproved. [37]

4. Conclusion and Outlook

In summary, it can be stated that the data acquired in the field of terminal block assembly can be considered as valid. This becomes particularly clear when looking at the distribution of frequencies in detail. This finding is an absolute basis for further analysis of the data set with regard to the influencing factors in manual assembly described above. This paper already provides important approaches for this. In addition to the identification of a physiological limit of the processing time, it could be shown that the influence of malfunctions in the operating process as well as incorrect operations increases with decreasing processing time. Finally, it was also shown by means of an example that a high degree of standardization is absolutely necessary in order to be able to calculate useful key figures on the basis of valid data. Furthermore, a small data basis means that the box plot method is less suitable for filtering the process values. When looking at the individual assembly orders, it has also become clear that a decreasing processing time per product with increasing work progress can be observed for the majority of the assembly lots recorded. This leads to the suggestion that there may be some form of learning. Here, a further assumption is that this effect occurs more strongly with increasing product complexity as well as with increasing batch size. In addition to this aspect, the influence of pause times and the influence of delivery deadlines will be the subject of further investigations. However, due to the complex correlations between the influencing factors, often only assumptions can be made. For a more precise analysis of the causes of human factors and for the creation of effect models, extended PDA systems and systematic experiments are required. The aim is also to vary and investigate influencing factors that were still set as fixed in the present use case. These include, among other things, the involvement of additional employees, the variation of the order sequence, and the modification of the process sequences, for example with regard to batch or piece production.

Acknowledgement

This research and development project is founded by the German Federal Ministry of Education and Research within the “SME – Innovative: Research for Production” Funding Action and implemented by the Project Management Agency Karlsruhe. The author is responsible for the content of this publication.

References

[1] Lotter, B., Wiendahl, H.-P. (Eds.), 2012. Montage in der industriellen Produktion: Ein Handbuch für die Praxis. Sammelwerk, 2. Auflage ed. Springer Vieweg, Berlin, Heidelberg.

[2] Wiendahl, H.-P., Gerst, D., Keunecke, L., 2013. Variantenbeherrschung in der Montage: Konzept und Praxis der flexiblen Produktionsendstufe (Einführung), Springer, Heidelberg

[3] Linsinger, M., 2021.: Situative Konfigurationsanpassung hybrider Montagesysteme, Dissertation.

[4] Sudhoff, M; Linsinger, M; Schulte, D; Kuhlenkötter, B. 2020. Development of a Concept for Real-Time Control of Manual Assembly Systems. In: Nyhuis, P.; Herberger, D.; Hübner, M. (Eds.): Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020). p. 92-98.

[5] Sudhoff, M., Leineweber, S., Linsinger, M., Niemeyer, J.F., Kuhlenkötter, B., 2020. Objective data acquisition as the basis of digitization in manual assembly systems, in: Procedia CIRP Vol. 93: 53nd CIRP Conference on Manufacturing Systems. p. 1176-1181.

[6] Kleppmann, W., 2013. Versuchsplanung: Produkte und Prozesse optimieren, 8., überarb. Aufl. ed. Hanser, München, 349 pp.

[7] DIN Deutsches Institut für Normung e. V., 2003. DIN 8593-1. Beuth Verlag GmbH, Berlin, 6 pp.

[8] Samy, S.N., ElMaraghy, H., 2010. A model for measuring products assembly complexity. International Journal of Computer Integrated Manufacturing 23 (11), 1015–1027.

[9] ElMaraghy, W. H., Urbanic, R. J., 2004. Assessment of manufacturing operational complexity. In: CIRP Annals – Manufacturing Technology, Band 53 (1). p. 401-406.

[10] Liu, P., Li, Z., 2012. Task complexity: A review and conceptualization framework. In: International Journal of Industrial Ergonomics, Band 42 (6). p. 553-568.

[11] Falck, A.C., Örtengren, R., Rosenqvist, M., Söderberg, R., 2017. Basic complexity criteria and their impact on manual assembly quality in actual production. In: International Journal of Industrial Ergonomics. 58. p. 117-128.

[12] Schlick, C., Bruder, R., Luczak, H., 2018. Arbeitswissenschaft, 4. Auflage ed. Springer Vieweg, Berlin, 740 pp.

[13] Schuh, G., Gartzen, T., Wagner, J., 2015. Complexity-oriented ramp-up of assembly systems. CIRP Journal of Manufacturing Science and Technology 10, 1–15.

[14] DIN Deutsches Institut für Normung e. V., 2016. DIN EN ISO 6385. Beuth Verlag GmbH, Berlin, 26 pp.

[15] Falck, A.C., Örtengren, R., Rosenqvist, M., Söderberg, R., 2016. Criteria for assessment of basic manual assembly complexity. In: Procedia CIRP, Band 44. p. 424-428.

[16] Wright, T.P. 1936. Factors Affecting the Cost of Airplanes. Journal of the Aeronautical Sciences 3 (4), 122–128.

[17] Greiff, M. de, 2001. Die Prognose von Lernkurven in der manuellen Montage unter besonderer Berücksichtigung der Lernkurven von Grundbewegungen. Zugl.: Duisburg, Univ.-Gesamthochsch., Diss., 2001, Als Ms. gedr ed. VDI-Verl., Düsseldorf, 158 pp.

[18] Biskup, D., 2001. Ablaufplanung mit gemeinsamen Due-Dates: Modelle, Lösungsverfahren und komplexitätstheoretische Klassifizierungen, Gabler Edition Wissenschaft ed. Deutscher Universitätsverlag, Wiesbaden, 189 pp.

[19] Liebau, H., 2002. Die Lernkurven-Methode. Ergonomia-Verl., Stuttgart, 191 pp.

[20] Landau, K. (Ed.), 1992. Die Arbeit im Dienstleistungsbetrieb: Grundzüge einer Arbeitswissenschaft der personenbezogenen Dienstleistung. Ulmer, Stuttgart, 505 pp.

[21] Maslow, A. H., 1943. A theory of human motivation. Psychological Review, 50(4), 370–396

[22] Fox, C., Levitin, A., Redman, T., 1994. The notion of data and its quality dimensions. Information Processing & Management 30 (1), 9–19.

[23] Ćwikła, G., 2013. Methods of Manufacturing Data Acquisition for Production Management - A Review. AMR 837, pp. 618–623.

24] Bokranz, R., Landau, K., 2021. Handbuch Industrial Engineering: Produktivitätsmanagement mit MTM, 2. Auflage ed. Schäffer-Poeschel, Stuttgart. 747 pp.

[25] Denkena, B., Dittrich, M.-A., Wilmsmeier, S., 2019. Automated production data feedback for adaptive work planning and production control. Procedia Manufacturing 28, 18–23.

[26] Hedderich, J., Sachs, L., 2018. Angewandte Statistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 1057 pp.

[27] Holland, H., Scharnbacher, K., 2004. Grundlagen statistischer Wahrscheinlichkeiten. Gabler Verlag, Wiesbaden.

[28] Scheid, S. C. 2001. Die verallgemeinerte Lognormalverteilung. Diplomarbeit, Dortmund.

[29] Berger, J., 2019. Wirtschaftliche Ungleichheit. Springer Fachmedien Wiesbaden, Wiesbaden, 362 pp.

[30] LIMPERT, E., STAHEL, W.A., ABBT, M., 2001. Log-normal Distributions across the Sciences: Keys and Clues. BioScience 51 (5), 341.

[31] O'Connor, P.D.T., Kleyner, A., 2011. Practical Reliability Engineering, 5th ed. ed. Wiley-Blackwell; John Wiley & Sons, Ltd, Oxford, 504 pp.

[32] Sachs, L., 1997. Angewandte Statistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 920 pp.

[33] Schuster, T., Liesen, A., 2017. Statistik für Wirtschaftswissenschaftler: Ein Lehr- und Übungsbuch für das Bachelor-Studium, 2. Auflage ed. Springer Gabler, Berlin, Heidelberg, 261 pp.

[34] Shrestha, J. 2019. P-Value: a true test of significance in agricultural research.

[35] Linsinger, M., Sudhoff, M., Lemmerz, K., Glogowski, P., Kuhlenkötter, B., 2018. Task-based Potential Analysis for Human-Robot Collaboration within Assembly Systems, in: Schüppstuhl, T., Tracht, K., Franke, J. (Eds.), Tagungsband des 3. Kongresses Montage Handhabung Industrie-roboter, Vol. 1. Springer Vieweg. pp. 1–12.

[36] Samy, S. N., ElMaraghy, H., 2012. A model for measuring complexity of automated and hybrid assembly systems. The International Journal of Advanced Manufacturing Technology, Springer-Verlag London. p. 813-833.

[37] Buck, J. R. 2013. Learning. In: Gass, S.I., Fu, M. (Eds.), 2013. Encyclopedia of operations research and management science, Third edition ed. Springer Science + Business Media, New York, NY.

Biography

Martin Sudhoff (*1991) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2017. He earned a master's degree in mechanical engineering at the Ruhr-University of Bochum. His primary research topics are the digitalization and automation of assembly systems.

Jonas Viehöver (*1994) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University of Bochum since 2020. He earned a master's degree in mechanical engineering at the Ruhr-University of Bochum. His primary research topic is automated quality management of assembly systems.

Michael Herzog (*1984) is a member of the Chair of production systems (LPS) at Ruhr-University Bochum as well as general manager of the interdisciplinary research center for the engineering of smart product-service systems (ZESS). His primary research topic is industrial servitization based on digitization, especially regarding the human-centred design of these systems.

Till 2009 **Bernd Kuhlenkötter** was responsible for product management and technology at ABB Robotics Germany. In 2009 Bernd Kuhlenkötter took over the Professorship for "Industrial Robotics and Production Automation" at the Technical University of Dortmund. Since 2015 he holds the professorship of "Production Systems" at the Ruhr-University of Bochum.

3rd Conference on Production Systems and Logistics

What Are The Role And Capabilities Of Fab Labs As A Contribution To A Resilient City? Insights From The Fab City Hamburg

Lennart Hildebrandt[1], Svenja Zadow[1], Luisa Lange[1], Michel Langhammer[1],
Manuel Moritz[1], Tobias Redlich[1], Jens P. Wulfsberg[1]
[1] Helmut-Schmidt-University / University of the Federal Armed Forces Hamburg

Abstract

Recent events such as the COVID-19 pandemic or the Ever-Given accident in the Suez Canal, which have led to local product shortages and negative social and environmental impacts, highlight the need to build resilience in areas that are highly affected by such events: in cities. One aspect of a multidisciplinary concept of resilient cities is the local manufacturing of physical products, which currently is mainly based on globally complex supply chains. The resilience of a city can be impaired if the supply of consumer goods can no longer be guaranteed, e.g., due to the fragility of supply chains. From this perspective, one of several pathways to a more resilient city is the emerging movement of open production sites (so called Fab Labs), where physical products can be produced or repaired in a distributed way by the consumers themselves. In metropolitan areas such as Hamburg, Fab Labs form networks including makerspaces, open workshops and educational institutions – so called Fab Cities. This article highlights the role of Fab Labs with regards to urban resilience and displays the capability of the Fab City Hamburg to contribute to the resilience of the city. To explore these capabilities in a pilot study, semi-structured interviews were conducted with makers and operators, and different Fab Labs were explored via participant observation. This article demonstrates that Fab Labs can contribute to a resilient city - especially from the perspective of manufacturing capability but also regarding the development of technical education. However, there are clear limitations with regards to the vertical range and manufacturing diversity.

Keywords

Resilient City; Urban Production; Production Sovereignty; Fab Lab; Fab City

1. Introduction

The COVID-19 pandemic has affected people around the world and, in some cases, drastically changed local living conditions. Contact restrictions and other limitations have been introduced to contain the spread of the virus. This has had a major impact on cities and their inhabitants, where, in many places, public life has been reduced to an absolute minimum and normal life was disrupted.

Despite these measures, very high numbers of COVID-19 cases occurred locally, leading to supply shortages of consumer goods, personal protective equipment and other devices (e.g., lung ventilators) especially early in the pandemic [1]. In the case of personal protective equipment one reason was that countries (specially in Asia) which manufacture personal protective equipment needed the equipment for their own needs at the start of the pandemic (e.g., China manufactures four-fifths of the world's protective clothing) [2,3]. However,

DOI: https://doi.org/10.15488/12142

it also became apparent that manufacturers outside the Asian region did not have the needed capacity and/or infrastructure to respond adequately to this high and short-term demand [4].

A similar lack of resilience through our global supply chains has occurred in the Ever-Given accident in the Suez Canal in March 2021, which had blocked the canal for six days. The problem was not so much the breakdown caused by the individual ship, but the six-day closure of one of the world's most important shipping channels. In total, 422 ships jammed and failed to reach their destination on time, causing delivery delays and damage to industry, wholesale and retail. This resulted in an estimated loss of supply of around 9 billion euros per day [5,6].

In both cases, entire countries were caught more or less unprepared due to the short lead time and immense and continuous demand for equipment. These problems manifest themselves especially where large crowds gather. In cities, there is a lot of contact with people in public spaces, a continuous and high demand for physical products, and at the same time, usually little manufacturing activity in the center. On the other hand, there are also large supply chains, but as the examples of COVID-19 and the Suez Canal show, they can be vulnerable to external impacts. These problems and resulting impacts on cities could even intensify in the future, as forecasts indicate that 70% of the global population will live in cities by 2050 [7].

For this reason, governments, scientists, urban planners and other actors have been coming up with concepts to make cities more independent (e.g. from global trade networks) and resistant to external influences (e.g. climate influences). The main contents of such concepts, which can be summarized under the term urban resilience, include new regional cooperations, new neighborhood developments and promoting civil society engagement and co-production.

Such approaches also exist in the Hamburg metropolitan region, where physical products have already been developed and manufactured by citizens. This approach, which is practiced in Hamburg and other cities globally, is called Fab City. In this paper, we present insights from the Fab City Hamburg and show what contribution Fab Labs can make to increase a city’s resilience in terms of production, innovation and education in order to be less dependent on external influencing factors.

2. Theoretical Background

2.1 Resilience and Resilient Cities

The concept of resilience is not a new notion and has its roots in the Latin word "resilire", which describes the ability of an entity or system to elastically restore its form and position after a disturbance or disruption. But as clear as its origins seem, as ambiguous seems its current definition. The polysemous concept of resilience varies in its meaning depending on the discipline, whereas this work focuses on urban resilience or resilient cities. The literature on this is numerous and has grown rapidly, especially in the last few years, with just about 40% of the total publications with the keyword "urban resilience" emerging in 2020 and 2021 [7]. For a detailed look at the definitions, see Merrow 2016 [8].

In their review of smart solutions and technologies during the pandemic, Sharifi et al. use the definition of the National Academies of Sciences, Engineering, and Medicine as a basis, as it combines different approaches and disciplines. According to this, resilience is defined as the capability to prepare and plan for, absorb, recover from, and more successfully adapt to actual or potential adverse events [9,7]. Sharifi et al. note that resilience is an approach to the management of socio-ecological systems that aims to develop an integrated framework to bring together the (often) fragmented, diverse research on disaster risk management [7]. In their review, Masnavi et al. [10] studied the exploration of the concept of urban resilience for use in urban planning and focused in detail at the conceptualization of a resilient city. As they explain, a common aspect of many approaches is the ability to withstand, resist and respond positively to pressure or change.

In addition, Patel et al. concluded, that resilience is not only about restoring functionality, but also about correcting existing social, political and economic structures that may have increased vulnerability and limited the ability to cope with the crisis [11]. It is also essential to note that resilience is not a rigid concept, but a dynamic set of conditions and processes [12].

The German Institute for Standardization has also published an ISO standard entitled "Sustainable cities and communities - Indicators for resilient cities" [13], which describes the concept of a resilient city in detail. Thus, by maintaining and restoring essential basic structures and services in a sustainable manner and through risk management practices, a resilient city is able to withstand, absorb, manage, adapt, transform and recover from the impacts of shocks and stresses [13].

In the meantime, numerous policy papers and reportings have been published, providing practical guidelines for making a city or region more resilient. Examples of this include the memorandum "Urban Resilience - Ways to create Robust, Adaptive and Viable Cities" of the German Federal Ministry of the Interior, for Building and the Home Affairs (BMI), which serves as a guideline for contemporary urban policy in Germany and Europe. It contains ten recommendations for action to increase a city's resilience, e.g., promoting civil society engagement and co-production or utilizing the potential of the neighborhood level [14]. Another example is the vision of the German Federal Environment Agency for a "City for Tomorrow", in which, for example, higher resource protection and more space for encounters are addressed [15].

A commonly shared aspect, when defining the term urban resilience or resilient city is an emphasis on the human factor and the contribution of human centered systems towards a city’s resilience, i.e. its ability and performance to anticipate, prepare, react, adapt and recover [16,17,12]. Based on these findings, the following working definitions of key terms of urban resilience is used for this paper:

- Anticipate - A city has the ability to foresee external changes.
- Prepare - A city has the ability to prepare for upcoming changes of external influences.
- React - A city has the ability to react at short notice to changes of external influences.
- Adapt - A city has the ability to adapt to changes of external influences in the long term.
- Recover - A city has the ability to recover from negative external influences.

2.2 Maker Movement, Fab City and Fab Labs

One approach to increase a city's resilience can be fostering technological literacy in conjunction with the creation of easy access production opportunities. This idea is part of the so-called Maker movement, originating from the USA, which focusses on self-production or community-based production of goods by prosumers (production by consumers) [18]. This new way of value creation offers potential in the areas of education, innovation and production [19]. In this movement, education (especially technical literacy) is fostered by an experimental learning environment, including rapid-prototyping methods of new product designs. In open-source communities, socio-technical contexts can be understood and evaluated and knowledge can be shared and accessed. This is a creative basis for innovation, as one's own problems can be better translated into technical and novel solutions through newly gained and existing knowledge. With this bottom-up approach, individual open source physical products (so called Open source hardware) can be developed and manufactured by their future users. “Open source hardware is hardware whose design is made publicly available so that anyone can study, modify, distribute, make, and sell the design or hardware based on that design” [20–22]. While the Maker Movement is more about fun, fabrication and community, essential content has been incorporated into the Fab City approach [23,24].

In the Fab City approach, founded in 2014, production is brought back to the city and carried out in the local neighborhood [25]. The goals of this city network (there are already 38 official Fab Cities distributed around the world) include greater independence from complicated transport logistics, sharing open (production) data in the network, building up or recovering technological literacy and developing new technologies at the point of need. The actual production in the neighborhood takes place in open production sites, so-called Fab Labs, with the help of easy-to-use digital production machines (e.g., 3D printers, CNC mills, laser cutters) [26]. Fab Labs are places where people can work creatively, innovate and produce physical products [27]. They provide access to production infrastructure for the local population through their equipment of hand tools and digital production machines. Besides individual products, small batches can also be produced, as exemplarily demonstrated by the production of face shields and masks during the COVID-19 pandemic outbreak in many places [28,29]. The term Fab Lab was first initiated at the Massachusetts Institute of Technology in 2002, and since then the number of such places has been growing rapidly: according to the Fab Foundation, a US non-profit organization for supporting the international Fab Lab network, there are already over 2000 such places worldwide.

Fab Cities with Fab Labs should be clearly differentiated from other innovative production approaches, which focus on raising the competitiveness on the market, such as frugal production or reconfigurable production systems [30,31]. The core strength of Fab Cities with Fab Labs is the triad of education, innovation and production as a whole. Fab Cities are more than innovative production approaches. Fab City "is about radical transformation, it is about rethinking and changing our relationship with the material world, in order to continue to flourish on this planet" [25].

Fab Labs and Fab Cities already provide an infrastructure that allows local innovation, urban production and knowledge building by everyone. However, full resilience of a city is clearly beyond the capabilities of a Fab Lab or current Fab Cities. But the need for a (more) resilient city is becoming increasingly clear. Therefore, the question arises: What contribution in terms of education, innovation and production can Fab Labs in Fab Cities make to achieve resilience in a city? In this paper, we present insights and results on the analysis of capabilities of the Fab City Hamburg based on selected Fab Labs.

3. Methodology

A qualitative and exploratory research approach with mixed-methods using data triangulation of semi-structured interviews and participant observation has been chosen for answering the research question. This approach was chosen since no data on this topic in cities is available in literature, therefore data was empirically collected as part of this study. According to R. K. Yin [32], a single context has been chosen. This is justified by the unique situation of the COVID-19 pandemic in combination with the Fab City approach, which is still not widespread, allowing data to be collected in a state that was previously inaccessible to scientists.

To explore the contribution of Fab Labs to increasing resilience, Hamburg was chosen as the research subject. Hamburg is home to more than 15 Fab Labs and other open production sites. Since 2019 the city has officially been part of the Fab City network. In the Fab City Hamburg, five universities and research institutions, three different authorities and various employees and volunteers are busy building a physical and digital infrastructure to be able to produce in the urban space of the city. Combined with the single context and the study of a single phenomenon (contribution of Fab Labs to increasing resilience) with multiple embedded and independent units of analysis (Fab Labs in Hamburg), this is an embedded single-case design.

Data was collected qualitatively in July 2021 by means of semi-structured interviews in nine Fab Labs with a total of 18 participants. For this purpose, an interview guideline was created with different questions that have an influence on resilience in the areas of education, production and innovation. These include internal lab processes, internal and external knowledge management, flexibility and speed of decision-making paths, demand messages, cooperation and collaborations, as well as production processes and planning and control systems in the respective Fab Labs.

The interviews were recorded in audio, anonymized, transcribed and evaluated and compared using qualitative content analysis according to P. A. E. Mayring [33]. This type of methodology was chosen in order to provide comparability between statements and information on the same topics through the semi-structured nature. At the same time, the semi-structured nature offers a great deal of freedom for interviewees to comment in greater depth on a subject, allowing the interviewer to gain new insights and internal knowledge.

One lab manager and one maker were interviewed per Lab. The group of people was intentionally limited and selected this way because we want to explore the influences on resilience and related events on site in a Fab Lab (shopfloor). The nine Fab Labs fall into three main categories: government/university Fab Labs, community-driven Fab Labs, and corporate Fab Labs. On the one hand this division was made to obtain the different views of a Maker or Lab manager for an individual Fab Lab. On the other hand, the categorization of the Fab Labs is necessary because significant factors with regards to people, technology and organization (e.g. financing, personnel, users, equipment, organizational structure and process organization, aim) differ significantly (for overview see figure 1). Since Fab City Hamburg continues to be in the making and only a few actors from the Maker Movement have contributed to this area so far, the interview partners from the three categories were chosen partly (2 Fab Labs) from the Fab City Hamburg environment, but mostly (7 Fab Labs) independently. However, the core statements of the interviews contained the same content, and a theoretical saturation was recognizable.

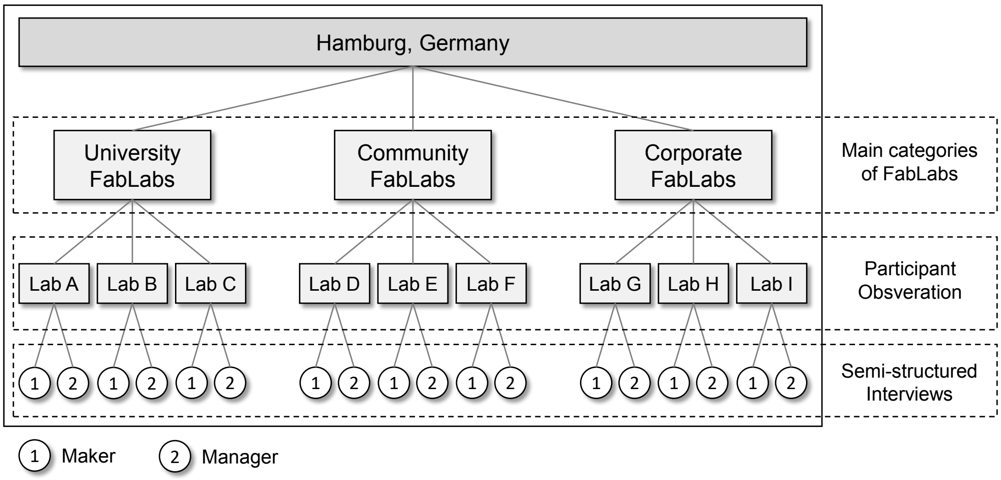


Figure 1: Research object

Additionally, as the second method of our mixed-methods approach, each Fab Lab was observed in a participatory (purely physical presence) and open manner. This observation was also conducted in a semi-structured manner using semi-structured observation guides to ensure comparability but also the possibility of deeper insights. The main focus of the observation was on the machines on site, which allowed to draw conclusions about the actual production capability. The results were systematized for the individual Fab Labs and the three main categories and documented in sub- and main tables. The evaluation was also carried out using qualitative content analysis. The mixed-methods approach combined with data triangulation and the large number of different Fab Labs and interview partners (in Hamburg) was chosen to best reduce potential bias.

4. Analysis and Results

In the following analysis, the findings and insights from the interviews and participant observation are used to evaluate the contribution of Fab Labs in the areas of innovation, production, and education promoting urban resilience. We see innovation, production and education as part of the solution to respond to the five identified fields of resilience. For this purpose, the definition derived in section 2.1 for building a resilient city, including the five sequences anticipation, preparation, reaction, adaption and recovering, is used to examine the contribution Fab Labs offer to building a resilient city in the dimensions of education, innovation and production (see figure 2). The processes behind each dimension are not considered as they are very individually dependent on the implementation of these dimensions by the actors. By deriving the dimensions through the already existing processes which were observed show the effectiveness and feasibility of the processes.

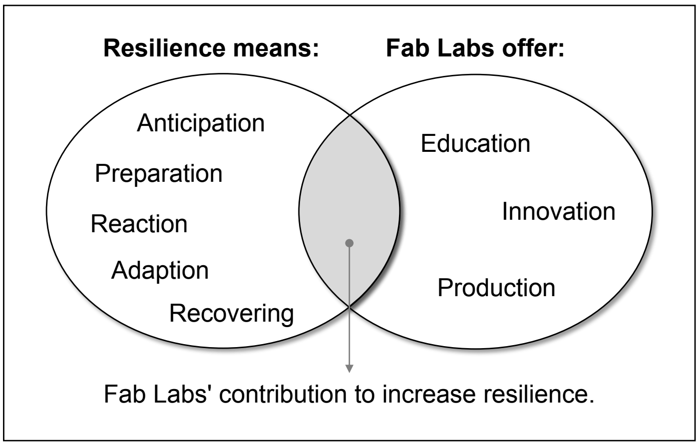


Figure 2: Intersection between resilience and Fab Labs

4.1 Education support

As the interviews show, the preparation of upcoming disturbances is supported by the Hamburg Fab Labs offering production spaces which are openly accessible, as well as offering openly organized workshops to transfer knowledge on how to use these spaces in terms of configuration of available machines and usage of computer aided design and development tools. These workshops reach from build workshops all the way to soldering, sewing and laser-cutting workshops. They take place at the labs directly and are organized for members of the local community as well as externals.

During the COVID-19 pandemic, it became clear that these workshops are essential to the use of Fab Labs, as they can uniformly bring a group of interested parties of six to twelve people to the point of being able to independently operate a machine as well as the associated software. For 3D printers, such a workshop lasts between two and four hours and can already be conducted with teenagers. Through these workshops as a preparatory measure, it was consequently possible to put people in a position to subsequently help as a workforce. In addition, participants then have the skills to use the site to develop and try out technical solutions, building further knowledge independently.

In addition to such workshops on the simple use of machines, there are also build workshops that are conducted throughout the community and usually rely on the use of open source hardware. The conducted build workshops represent the process of local replication of open source hardware designs. This openness enables a deeper insight into the technical fundamentals of the hardware as such and provides opportunities of re-designing as well as repairing the hardware. The workshop offers are mostly bundled and available on one particular day of the week (so-called open lab day). Through these open offers and the collaborative organization of these workshops by the community and managers, knowledge is passed on to the civilians, technical literacy is built up and promoted, and the resilience of a city is fostered through the possibility of self-production of consumer products and machines.

Interviews with the lab managers, particularly in community-driven fab labs, have shown that the lack of professional staff to organize and conduct these workshops and the financing of the activities are the most common issues hereby. This problem is much less prevalent in corporate Fab Labs, where personnel capacities for workshops are more limited (due to lack of time). Another problem is the low level of cooperation between the Fab Labs. There is no Fab Lab in Hamburg that has a strategic cooperation with other Fab Labs, which means that resources (including resources for workshops) can only be used inefficiently.

"Unfortunately, there is no real cooperation. It's just a bit, when we are overloaded, I send our students somewhere else to print. But cooperations unfortunately not." - *University Fab Lab manager*

The constant documentation of this knowledge is essential for the use of open knowledge and the (machines building) workshops already described. Accompanied by the offer of open educational resources provided by Fab Labs in Hamburg, the knowledge is thus digitally available and regardless of location. Lessons learned throughout shock periods are accumulated in the described open-ended manner and can be used for new learning cycles. Thus, the skills learned are retained and provide a long-term basis for responding to new events. Such concepts are essential to form an educational basis to build up preparational and recoverable resilient structures in cities. As the interviews have shown, the documentation of this knowledge is still a widespread problem, since the existing knowledge is insufficiently processed and summarized for posterity, which currently still leads to difficult learning conditions.

4.2 Innovation support

Fab Labs offer a significant contribution to fostering innovation, through their open workshop, whereby users have the opportunity to develop technical solutions to problems and produce prototypes. This gives rise to new products that can subsequently be used in the local area. Users are supported in their innovations by innovation consultants or by the lab managers, who are available to answer questions about machines and technical solutions during the respective opening hours (open lab days). However, it is also apparent that insufficient funding and a lack of knowledge mean that the potential of this advice cannot yet be fully exploited.

Fab labs in conjunction with open source hardware offer great potential for innovation, as the designs of these products are available digitally and may be adapted by any user. The knowledge of a larger community lead to quick design adjustments and improved products. Fab Labs also offer a way to quickly develop and try out new prototypes (similar to the lean startup approach of "build, measure, learn"). According to the interviews, in Hamburg this knowledge and opportunity was massively used during the COVID-19 pandemic, when Face Shields and Face Masks based on an open source hardware design were developed, refined, and produced in several Fab Labs (especially base-funded university Fab Labs). This rapid reaction enabled the products to be adapted to local (especially climatic and regulatory) constraints and subsequently, in conjunction with the manufacturing infrastructure, to supply several hospitals in Hamburg, Schleswig-Holstein and Madrid. [28,29]

"A colleague then presented the first adapted concepts for the face shield, for example. Then we went into the first prototype phase, where we tried to see what could be printed and how, and how the whole process could be optimized. Then again a bit of try and error. First attempts at how we can stack things for maximum effectiveness. And then it worked." - *University Fab Lab user*

The innovation potential described above can also be found in the area of adaptation and recovering, where technical solutions can be (further) developed to meet local needs. An example from an interview with a manager of a university Fab Lab shows the potential of local communities in idea competitions, where the citizens are invited to contribute to specific topics where new innovative solutions are needed, or existing solutions are optimized or redesigned based on local gathered requirements.

This community-based innovation process operated and curated in Fab Labs supports the city's ability to prepare for upcoming events by designing new technical solutions, to react to events by deploying the solutions in a rapid way and to adapt flexible to new states of urban settings. In doing so, the open innovation cycle is iterative and always adapts to new circumstances, which can also go beyond a state of shock.

4.3 Production support

The creation of a responsive and adaptable city infrastructure is supported by Fab Labs through the provision of small open physical production facilities and machines. This production infrastructure provides a lever for action to produce needed products locally and as far as possible without dependencies in the global supply chain. This allows us to build and maintain a production sovereignty that enables us to manufacture physical (simple) everyday objects. The big advantage here is the low variance of the raw materials: for example, the single raw material (filament) of the 3D printer can be used to produce a very wide variety of products.

The machines in a Fab Lab vary but are mostly standardized. The most important machines as well as the materials to be machined, typical machining times and typical usage of the machines (based on the interviews) are shown in table 1. Fab Labs cover a large scale of plastic processing tools (especially through the use of 3D printers). In addition, metal and wood are other materials that can be processed locally. However, the interviews show that there are limitations in the processing complexity. The given degrees of freedom in processing are typically limited to 3 axes (x, y, z direction) in the examined machines (for example 3-axis CNC milling machines). In terms of the value of the machines, it has been shown that university and corporate Fab Labs tend to have more financial resources to purchase higher quality and multiple machines related to a manufacturing process. In community-driven Fab Labs, on the other hand, there are many cheaper machines or self-built solutions (as open source hardware) that are, however, very well adapted and appropriate to their area of application.

Table 1: Machines in Fab Labs

Digital machines	Materials	Processing time	Dissemination	Usage
3D printers	Plastics (especially PLA, TPU, ABS)	Hours to days	Very high	High
Laser cutters	Especially wood, plastics, metal	Minutes	High	Very high (most used machine)
CNC mills and lathes	Wood, plastics, partly metal	Minutes, partly hours	Medium	Medium
Vinyl cutters	Vinyl, paper	Minutes	Low	Low
Circuit board printer	Circuit boards	Minutes	Low	Medium
3D scanners	Any	Minutes	Low	Medium

Due to these restrictions, there are clear limitations with regards to the vertical range of manufacture, which is additionally restricted by the quality of the machining. However, these machines are perfectly adequate for the production of physical products where lower quality requirements are placed on the machining and/or only simpler geometries are used. As a current example, the production of officially approved face shields in larger quantities (up to 5,000 face shields per week) with 3D printers and laser cutters in a single Fab Lab should be mentioned [28,29]. In addition to the manufacturing of physical products, the existing machines and tools can also be used for the maintenance and repair of goods which support the city's resilience recovery process.

With regard to the possible output of a Hamburg Fab Lab network (i.e. Fab City Hamburg) and the ability to react to external influences, the potential of a distributed and networked production infrastructure with Fab Labs was also analyzed. The interview results show that Fab Labs as locations see themselves as independent organizations with mostly no manufacturing organization.

> *"There is almost no production organization structure at all. [...] For example, someone in his telegram group throws something into the room about what needs to be done or an idea about what we can do, and then we talk about it. This is then organized somehow." - Community Fab Lab manager*

Fab Lab managers and Makers are also primarily focused on a single Lab. This leads to manufacturing that is currently not very efficient, since, as the interviews show, machine capacities are unnecessarily kept on hand and/or not used. However, the individual Fab Labs cooperate regionally with actors such as universities, schools, SMEs or associations in a cooperation network, but as shown not in a production network.

In addition to the physical infrastructure, some of the Hamburg Fab Labs (especially university and community driven Fab Labs) host and use open-source software tools for design and development of open-source hardware (e.g., image processing programs or CAD/CAM software) or for operating their organization (e.g., open source cloud software and development platforms with built-in version control). Through this use, Fab Labs can offer digital infrastructure for production systems which stays digital sovereign and flexible used by community members and external partners to create and exchange knowledge. In terms of the five resilience sequences, this offers advantages particularly in preparation, reaction and adaptation, since the software infrastructure cannot be easily deactivated from the outside, e.g. in a defense case. It also ensures that the knowledge (e.g. designs, production data) stored on the local servers and in the local software applications is available for the long term. The interviews showed that such a data structure is currently under development, especially in the university and community-driven Fab Labs.

Besides the production capability, the factor of decision-making process in Fab Labs is important for the flexible reaction and adaptation in a resilient city. This process depends on the legal and organizational structure. The observed community driven Fab Labs in Hamburg are operated as associations. Decisions are made by consensus at the board or member level. For larger and strategic decisions, there is usually a two- to four week lead time to vote on issues that need to be decided. This circumstance is an obstacle to flexibility, though it is a very democratic process. In commercial and university driven labs, decisions are made by the responsible lab managers or delegated to corresponding employees. The consistently small-scale organizational structures of these Fab Labs enable fast and flexible decision-making, which contributes to an optimized implementation process when needed (see face shield and face mask production during the COVID-19 pandemic in [28,29]. In general the effort of processes is very dependent on the organizational structure of the actors.

After providing an overview of Fab Labs' contributions to resilience enhancement, the following section discusses key findings and identifies limitations. On this basis, the authors recommend a research agenda for future in-depth investigations

5. Discussion and Research Agenda

The aim of this paper was to elaborate in a pilot-study the role and capabilities of Hamburg Fab Labs as a contribution to making Hamburg a resilient city. As displayed in the analysis, the production capacity and product variety of the Hamburg Fab Labs is still very limited. Currently, only the processing of simple geometric products and/or products with lower quality and quantity requirements is possible. Furthermore, especially the community driven labs are currently facing the major issue of personnel shortage. This is partially due to the difficulty of funding, which is why the work is largely based on volunteer labor. University and corporate driven labs do not/less have the problem.

Even though their potential is currently greatly reduced e.g., by a lack of cooperation among the individual labs, especially the regional cooperations deriving from singular Fab Labs can still have a larger impact on strengthening a city as such [34]. As seen in the analysis, the Fab Labs in Hamburg have several ways of contributing towards the resilience of the city of Hamburg. Through their flexible adaptation and quick reaction, they play a significant role when it comes to Hamburg's ability and performance to prepare, react, adapt and recover. Anticipating structures, however, are not covered by Fab Labs as they do not have capabilities of foreseeing external disturbing events or shocks. Especially within the area of material processing, Fabs Labs in general prove as great support with regards towards a city's resilience. Additionally, the innovational and educational contribution of the Fab Labs are important factors when it comes to strengthening urban resilience.

It must be kept in mind that resilience is a concept which is rather abstract. There can never be "the resilient city" [35] and furthermore, a city will never become resilient by the mere existence of Fab Labs and the Maker Movement. Nevertheless, it can be noted that Fab Labs in general can contribute to strengthening a city's resilience to certain events, such as shortages of certain products or natural disasters, by keeping a production infrastructure in the neighborhood, by sharing knowledge, and as places of rapid and adaptive innovation. As general key potentials of Fab Labs are the areas of education, innovation and production [19], the main findings (both advantages and disadvantages) of this study can be applied to other Fab Labs as well. However, since this study only refers to selected Labs in the Hamburg area, it is recommended to take a closer look at other facilities. A comparative study that includes additional German and European cities would underpin this concept. A focus can be placed on comparing regions with and without a Fab Lab network or a strong Maker movement.

However, this pilot study is also just the start of data collection. To more deeply understand the Fab Labs' contribution and potential, we specifically propose the following research agenda, first for the Hamburg area (for deeper understanding), and subsequently in further national and international (Fab) cities with a wider range of interview partners (both in terms of number and in terms of orientation, e.g. adding municipal partners or corporate partners):

1. **Product Design & Production**:
   How can easier-to-use and more digitized (open source) machines enable Fab Labs to achieve a greater design complexity, greater production depth, and greater output?
   Which effect has an improved strategic collaboration and networked production planning and control systems on the output of Fab Labs?
2. **Education and interactions**:
   How do Fab Labs influence local neighborhoods and the technological literacy of users?
   What further potential do Fab Labs bring within these areas?
3. **Finance**:
   Which new funding models for staff, particularly in community Fab Labs, that support on-the-ground work, are suggested?
   Which further limitations within the labs' volunteer work occur?
   Which types of financing models are recommended for Fab Labs, allowing short-term and flexible use without membership or other obligations?
4. **Further sectors**:
   Which influence do Fab Labs have on further fundamental sectors such as healthcare, waste management or food?

6. Conclusion

This paper evaluates and reflects on the role and capability of Fab Labs in the Fab City Hamburg as a contribution to a resilient city. In order to investigate this concept in a pilot study, semi-structured interviews and participant observations have been conducted in nine different Fab Labs within city. The results of these interviews and observations have been compared with the current state of research towards creating a resilient city. For this purpose, a working definition of a city's resilience, including the ability and performance to anticipate, prepare, react, adapt and recover, was defined and cross-referenced with the current contributions of local Fab Labs in the areas of education, innovation, and manufacturing. The data and results presented in this paper suggest that Fab Labs can have a strong contribution towards a city's resilience. In particular, the ability to build technical knowledge locally in the neighborhood, to develop technical innovations alone or in groups, and to subsequently manufacture prototypes or physical products are main drivers towards forming a resilient city, which are strengthened and enhanced throughout Fab Labs. These drivers form the core of the Fab Labs` contribution to support cities in their approach to resilience. However, this study also highlights the limitations of Fab Labs for example in the manufacturing context with regards to manufacturing depth and diversity. Future work within this area should include studies focusing on other (inter-)national cities with a wider range of interview partners and different stakeholders. Essential will be further research on production machines and networks, funding for human resources and Fab labs, and influences on education and neighborhoods to more deeply understand the contribution of Fab Labs to urban resilience.

7. Acknowledgements

This research paper is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr which we gratefully acknowledge [project Fab City].

8. References

[1] Pearce, J.M., 2020. A review of open source ventilators for COVID-19 and future pandemics. F1000Research 9, 218.

[2] Buckley, C., Wee, S.-L., Qin, A., 2020. China's Doctors, Fighting the Coronavirus, Beg for Masks. nytimes.com.

[3] Grill, M., Kampf, L., 2020. Merkel macht's möglich. Atemschutzmasken aus China. tagesschau.de.

[4] Giesen, C., Ott, K., Richter, N., 2020. "Direkter Zugang" zu Schutzkleidung aus China. sueddeutsche.de.

[5] Clark, A., 2021. Suez Snarl Seen Halting $9.6 Billion a Day of Ship Traffic. bloomberg.com.

[6] Hosseini, S., Ivanov, D., Dolgui, A., 2019. Review of quantitative methods for supply chain resilience analysis. Transportation research / E.

[7] Sharifi, A., Khavarian-Garmsir, A.R., Kummitha, R.K.R., 2021. Contributions of Smart City Solutions and Technologies to Resilience against the COVID-19 Pandemic: A Literature Review. Sustainability 13 (14), 8018.

[8] Meerow, S., Newell, J.P., Stults, M., 2016. Defining urban resilience: A review. Landscape and Urban Planning 147, 38–49.

[9] Cutter, S.L., Ahearn, J.A., Amadei, B., Crawford, P., Eide, E.A., Galloway, G.E., Goodchild, M.F., Kunreuther, H.C., Li-Vollmer, M., Schoch-Spana, M., Scrimshaw, S.C., Stanley, E.M., Whitney, G., Zoback, M.L., 2013. Disaster Resilience: A National Imperative. Environment: Science and Policy for Sustainable Development 55 (2), 25–29.

[10] Masnavi, M.R., Gharai, F., Hajibandeh, M., 2019. Exploring urban resilience thinking for its application in urban planning: a review of literature. Int. J. Environ. Sci. Technol. 16 (1), 567–582.

[11] Patel, V., Sharma, A., Lal, R., Al-Dhabi, N.A., Madamwar, D., 2016. Response and resilience of soil microbial communities inhabiting in edible oil stress/contamination from industrial estates. BMC Microbiol 16 (1), 50.

[12] Turnbull, M., Sterrett, C.L., Hilleboe, A., 2013. Toward resilience: A guide to disaster risk reduction and climate change adaptation. Practical Action Publishing, Warwickshire.

[13] ISO 37123 - 2019-12 - Beuth.de, 2022. https://www.beuth.de/de/norm/iso-37123/317795700. Accessed 27 January 2022.

[14] Bundesministerium des Innern, für Bau und Heimat (Ed.) 2021. Memorandum Urbane Resilienz: Wege zur robusten, adaptiven und zukunftsfähigen Stadt, 17 pp. Accessed 20 January 2022.

[15] Umweltbundesamt, 2022. Die Stadt für Morgen: Die Vision. https://www.umweltbundesamt.de/themen/verkehr-laerm/nachhaltige-mobilitaet/die-stadt-fuer-morgen-die-vision#zusammenleben. Accessed 27 January 2022.

[16] Arafah, Y., Winarso, H., Suroso, D.S.A., 2018. Towards Smart and Resilient City: A Conceptual Model. IOP Conf. Ser.: Earth Environ. Sci. 158 (1), 12045.

[17] Homagk, L.-M., 2019. Urbane Resilienz – Ein brauchbares Konzept für die Steuerung der Stadtentwicklung? : Erfahrungen und strategische Empfehlungen am Beispiel der Stadt Hamburg : experiences and strategic recommendations based on the example of the city of Hamburg.

[18] Redlich, T., Wulfsberg, J.P., Bruhns, F., 2008. Virtual Factory for Customized Open Production. Proceedings of the 15th International Product Development Management Conference.

[19] Hildebrandt, L., Moritz, M., Seidel, B., Redlich, T., Wulfsberg, J.P., 2020. Urbane Mikrofabriken für die hybride Produktion. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (4), 191–195.

[20] https://www.oshwa.org/definition/, 2022.

[21] Moritz, M., Redlich, T., Grames, P.P., Wulfsberg, J.P., 2016 - 2016. Value creation in open-source hardware communities: Case study of Open Source Ecology, in: 2016 Portland International Conference on Management of Engineering and Technology (PICMET). 2016 Portland International Conference on Management of Engineering and Technology (PICMET), Honolulu, HI, USA. 9/4/2016 - 9/8/2016. IEEE, pp. 2368–2375.

[22] Moritz, M., Redlich, T., Wulfsberg, J., 2018. Best Practices and Pitfalls in Open Source Hardware, in: Rocha, Á., Guarda, T. (Eds.), Proceedings of the International Conference on Information Technology and Systems (ICITS 2018), vol. 721. Springer, Cham, pp. 200–210.

[23] Hatch, M., 2014. The maker movement manifesto: Rules for innovation in the new world of crafters, hackers, and tinkerers. McGraw-Hill Education, New York.

[24] Ramsauer, C., Friessnig, M., 2016. Einfluss der Maker Movement auf die Forschung und Entwicklung, in: Biedermann, H. (Ed.), Industrial Engineering und Management. Beiträge des Techno-Ökonomie-Forums der TU Austria, 1. Aufl. 2016 ed. Springer Gabler, Wiesbaden, pp. 43–61.

[25] Diez, T., 2016. Fab City Whitepaper: Locally productive, globally connected self-sufficient cities.

[26] Rumpala, Y., 2021. 'Smart' in another way: the potential of the Fab City approach to reconfigure urban dynamics. Urban Research & Practice, 1–23.

[27] Sheridan, K., Halverson, E.R., Litts, B., Brahms, L., Jacobs-Priebe, L., Owens, T., 2014. Learning in the Making: A Comparative Case Study of Three Makerspaces. Harvard Educational Review 84 (4), 505–531.

[28] Hartig, S., Duda, S., Hildebrandt, L., 2020. Urgent need hybrid production - what COVID-19 can teach us about dislocated production through 3d-printing and the maker scene. 3D printing in medicine 6 (1), 37.

[29] Hildebrandt, L., Redlich, T., Wulfsberg, J.P., 2020. Persönliche Schutzausrüstung aus der hybriden urbanen Mikrofabrik. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (9), 576–580.

[30] Boldt, S., Rösiö, C., Bergström, A., Jödicke, L., 2021. Assessment of Reconfigurability Level within Existing Manufacturing Systems. Procedia CIRP 104, 1458–1463.

[31] Herstatt, C., Tiwari, R., 2015. Frugale Innovation. WIST 44 (11), 649–652.

[32] Yin, R.K., 2014. Case study research: Design and methods, 5. edition ed. SAGE, Los Angeles, London, New Delhi, Singapore, Washington, DC, 282 pp.

[33] Mayring, P., 2010. Qualitative Inhaltsanalyse: Grundlagen und Techniken, 11., aktual. und überarb. Aufl. ed. Beltz, Weinheim, Basel, 144 pp.

[34] Hildebrandt, L., Redlich, T., Wulfsberg, J.P., 2021. Production Planning And Control In Distributed And Networked Open Production Sites – An Integrative Literature Review. Proceedings of the 2nd Conference on Production Systems and Logistics (CPSL 2021).

[35] 2018. Stresstest Stadt - wie resilient sind unsere Städte?: Unsicherheiten der Stadtentwicklung identifizieren, analysieren und bewerten. Bundesinstitut für Bau-, Stadt und Raumforschung.

9. Biography

Lennart Hildebrandt (*1994), M.Sc. studied mechanical engineering (product development and logistics) at Helmut-Schmidt-University in Hamburg and joined the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering there in 2019. His research focus is on value creation systematics and urban production in connection with Fab City and open source hardware.

Svenja Zadow (*1990), M.Sc. studied agricultural economics at Kiel University (CAU) and joined the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in 2021. Her research focuses on value creation systematics, sustainable consumption, urban food production and occurs in the context of Fab City.

Luisa Lange (*1993), M.Sc. studied mechanical Engineering at UAS Constance and Jönköping University in Sweden. She joined the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in 2021. Her research focuses on value creation systematics and the human factors engineering in new connection with Fab City and open source hardware.

Michel Langhammer (*1990), M.Sc. studied electrical engineering (automation systems) at Baden-Wuerttemberg Cooperative State University (DHBW) Stuttgart and Hamburg University of Applied Sciences (HAW). He joined the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in 2021. His research focuses on open source hardware, sustainable manufacturing and open educational resources.

Dipl-Ing. Manuel Moritz (*1985), MBA studied mechanical engineering at Helmut-Schmidt-University in Hamburg and joined the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering there in 2015. His research focus is on value creation systematics in connection with Fab City and open source hardware.

Dr.-Ing. Dipl.-Wirt.-Ing. Tobias Redlich (*1981), MBA is the senior engineer at the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in Hamburg since 2013. His research focuses on value creation systematics and the impact of new value creation patterns on production engineering and production management.

Univ.-Prof. Dr.-Ing. Jens Wulfsberg (*1959) is head of Chair of Production Engineering in the Department of Mechanical Engineering at Helmut-Schmidt-University in Hamburg since 2001. The main research areas at the Manufacturing Technology Laboratory (LaFT) are value creation systems, robotics, micro manufacturing as well as additive manufacturing and lightweight construction.

3rd Conference on Production Systems and Logistics

Modular Software Architecture For Interfacing Online Scheduling Agents With Assembly Planning And Control Systems

Amon Göppert[1], Jonas Rachner[1], Lea Kaven[1], Robert H. Schmitt[1,2]
[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

Production systems must be more resilient and adaptive due to mass customization and increasingly external disturbances, such as supply chain disruptions or changing policies. As the last chain in the production value stream, assembly systems are especially prone to fluctuations, leading to alternative and more flexible assembly system designs. Online scheduling is a crucial component for dynamically controlling a flexible assembly system.

This work presents a modular software architecture that interfaces between online scheduling agents and control systems. A standardized data model of the assembly system allows for exchanging different scheduling agents during the planning or operation phase. Applications are benchmarking competing algorithms, validating scheduling results by comparison, and seamlessly substituting or updating scheduling algorithms. The standardized data model and interface on the assembly system side facilitate the transition between planning and operation. A simulation model can be interchanged with a control system without extra effort to integrate the control system's scheduling agents. Additionally, the modular architecture enables production-parallel simulation to optimize the running system by evaluating and executing alternative scenarios.

The long-term assembly system performance can profit from the modular architecture by updating the agent during production with advances in online scheduling algorithms (e.g., machine learning). Furthermore, the modular architecture enables the required resilience and adaptability by fast switching from simulation to real control systems and supporting system optimizations during operation.

Keywords

Production control; software architecture; adaptive assembly; online scheduling; simulation

1. Motivation

Various global trends in production technology push producing companies towards being more resilient and adaptive: Customer’s demands are increasingly volatile and individualization is an ongoing change driver. [1,2] Additionally, the recent pandemic still disrupts globalized supply chains and causes a shortage in semiconductors for production systems. [3] As a consequence, production systems and especially assembly systems, as the vulnerable last piece in the production chain, need to be adaptive and resilient to changes. [4,5] A solution are alternative assembly systems that break with the traditional concept of takt time and linear transfer in classical assembly lines. Examples of such concepts are matrix assembly systems [6,7], modular assembly systems [8], agile assembly systems [9], or dynamically interconnected assembly systems [10]. The break-up of linear transfer in those flexible assembly systems allows job routes that can be determined individually for each job. The allocation of processes at work stations for the jobs along the job

DOI: https://doi.org/10.15488/12134

route can occur dynamically and with short reactions times. A consequence of the greater flexibility with job routes and the more dynamical character is a greater complexity in the landscape of software components for planning and control of such assembly systems. Also, in those systems changes in the product mix or disturbances, due to rework, missing components or worker shortages, are occurring in high frequency. [11,12] Therefore, online scheduling is a relevant task in such assembly systems. Machining systems are not considered, due to the significantly longer process time, less disturbances, and consequently less need for frequent re-scheduling or online scheduling.

Following the complexity and the demands from online scheduling, multiple specific challenges for interfacing the different software components arise: During production in dynamical assembly systems with many disturbances, online scheduling algorithms must ensure the efficient job allocation. [13] These algorithms must cope with the complexity and temporal variability of boundary conditions. Therefore, traditional rule-based heuristics are not applicable and machine learning (ML) algorithms need to be applied. [14] These machine learning algorithms for online scheduling need to be connected to the control system that requests solutions for job allocation problems. The machine learning models are trained with simulation models that enable the generation of large datasets, required for machine learning training. Due to the significant computational time needed for most machine learning models they are trained before production [14] or with delay until deployment during production [15]. Consequently, in both cases an interface needs to enable the communication between the required simulation models and the online scheduling algorithms. Additionally, during training, a benchmark of different algorithmic approaches supports the choice over an algorithm. During production, alternative parametrizations or algorithms can be tested and deployed for the application with the control system. This demand for exchangeable online scheduling algorithms also results in an interface between the algorithms and the control system or simulation model.

A possible solution for these complexity-driven challenges could be a modular and standardized software architecture that seamlessly connects the online scheduling algorithms with the simulation and control components during planning and production. In the next chapter 2 the state of the art in scheduling and control architectures, the derived research question and in brief the applied methodology are described.

2. State of the Art and Methodology

Literature presenting research regarding online scheduling primarily concentrates on the performance improvement or evaluation of scheduling algorithms, for instance, with the application of deep reinforcement learning techniques, but not on the integration in dynamical assembly system planning and control. [16–19] The existing literature on architectures for production planning and control for flexible production systems focuses on adaptive architectures as multi-agent systems in different heterarchical or hierarchical structures [20–22] or cloud manufacturing paradigms. [23,24] In those architectures, an agent is defined as a computer system embedded in an environment, which has the ability to perform autonomous actions to achieve predefined goals. [25] Online scheduling algorithms are included as agents connected to the shopfloor environment (e.g. [26]). Other potentially suitable architectures present modular control systems that incorporate approaches for online scheduling. [27] The presented control architectures typically focus on integrating online scheduling algorithms in the production phase and not on the application of the agents in the planning phase through connection to simulation environments. However, as explained above, simulations are indispensable e.g. for training ML algorithms. Therefore, a large body of literature integrating online scheduling algorithms into simulations exists (e.g. [28,29]), but neglects the integration of the proposed algorithms into existing control systems.

As a summary, existing architectures contain agent-based approaches that enable a modularity for the scheduling agents and they provide interfaces for the superordinate planning or control systems. But, the reviewed architectural approaches either focus on the application of online scheduling in connection with simulation models that run independently from real production systems. Or they only focus on the application in multi-agent systems that are connected to the shopfloor during production. Therefore, the deficit can be concluded that online scheduling agents cannot be connected to simulation models during planning or ML training and not to control systems in a multi-agent system approach during production with the same interface in current architectures. In other words, in the analysed literature, no dedicated software component for independently managing online scheduling requests and decisions. The following research question can be derived from this deficit:

> *How can the seamless connection of various online scheduling agents with planning and control systems be enabled?*

To answer this research question the developments focus on newly overall new architecture, the required components and the necessary standardized interfaces. The well-established multi-agent system paradigms, including modular scheduling agents [20–22], already existing interface data models for transferring data in online scheduling scenarios [27–29] and principles of service-oriented architectures [30,31] are incorporated and used for the new architecture concept.

The following chapter 3 presents the modular software architecture to address the above-stated deficit on a conceptual level. Two representations are used for this. First, a software architecture oriented overview on the components and their interfaces is given and described in detail. Second, an incorporation of the components in the layers of the Internet of Production reference framework is presented to provide a different perspective from production technology. The subsequent chapter 4 presents an implementation and testing of the conceptual architecture. The overarching structure of chapters 3 and 4 is based on the methodology of software and model development with the phased of formalization (conceptual model), implementation (executable model) and experimentation (testing results). [32]

3. Conceptual Modular Software Architecture

Figure 1 presents an overview of the modular software architecture. The proposed architecture is divided into two phases and two layers. In the system layer, the simulation module and the control system represent the digital version of the assembly shopfloor. The simulation module is applied to train machine learning models for online scheduling. Furthermore, independently of online scheduling, the simulation can evaluate assembly system alternatives, which is applied mostly during planning, but also during production to optimize the system. The control system, or manufacturing execution system, is responsible in the production phase to monitor and execute assembly or auxiliary processes at work stations, buffers, or the transport system. In a flexible alternative flexible assembly system as explained in the motivation, a job route needs to be scheduled during production respectively during simulation run time. After a process step, the job might have a variety of next process-station combinations as potential alternatives.

The simulation or control system gathers the required information in a data model for decision-making. This data model needs to comprise all relevant static parameters describing the overall system configuration and dynamic variables representing the current system state. The parameters and variables need to fully represent the three system categories products (e.g. jobs allocated at processes and work stations), processes (e.g. progress, sequence, durations) and resources (e.g. transportation system, buffers, work stations states). The data model is standardized to allow interoperability. The ontology-based definition enables a standardized meta-model of digital twins as a connected data model. (cf. [33,34]) The interface methods are implemented continuously through the planning or production phase, i.e. in simulation and control systems. The methods are responsible for connecting to the scheduling layer. A scheduling request sends the data model of the assembly system.

The scheduling server, as a crucial part of the architecture, builds the interface between the system layer and the scheduling agents. It processes the scheduling request and forwards it with interfacing methods to one or multiple scheduling agents. Due to their independent decision-making and potential ability to learn in exchange with the environment, the term scheduling agent is chosen.

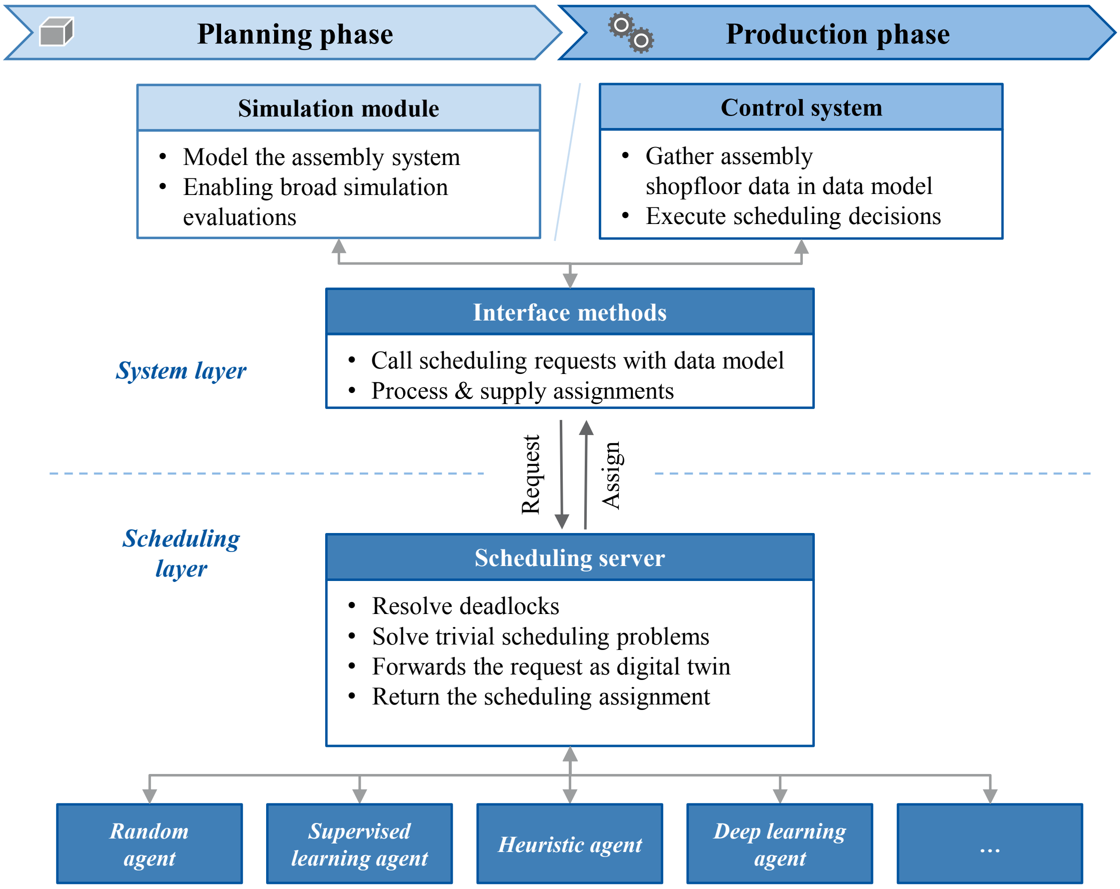


Figure 1: Overview of the modular software architecture for online scheduling.

In addition to communication functionalities, the scheduling server abstracts repetitive and simple tasks from the scheduling agents to reduce time-to-response. In case of trivial scheduling problems with only one station-process combination, the scheduling server can directly respond to the system layer with a scheduling assignment without addressing the scheduling agents. Another encapsulated task in the scheduling server is deadlock resolution. Deadlocks in online scheduling occur, for instance, when two jobs cross request the current station of the respective other job. In such cases, the scheduling server must force a scheduling decision or send a job to an intermediate buffer to free a station. As deadlock resolution is not a specific scheduling problem, but more a general production control issue, the abstraction avoids unnecessary communication and duplicate implementation in the scheduling agents.

The decision-making to provide a scheduling assignment occurs in the scheduling agents. As described in the motivation, various scheduling agents can be deployed: A random agent as a random performance baseline for algorithm comparison, a heuristic agent with rule-based logic, or agents applying machine learning techniques such as deep reinforcement, unsupervised or supervised learning. The type of machine learning technique is independent for the scheduling server as it communicates via the request and assignment messages. The scheduling agents can be seen as suppliers of scheduling-as-a-service in the modular architecture.

The scheduling decision is a process-station combination. The scheduling agent returns it to the scheduling server, which replies to the interface methods that supply the results to the simulation or control system. There, the results are used to execute simulated or real processes or actions. For instance, during production, the scheduling decision triggers a handling unit and a transport system to move the job to the next work station.

As an alternative representation, Figure 2 shows the modular software components from Figure 1 in the layers of the Internet of Production framework. [35] In comparison with Figure 1, this representation aims at the allocation of components in an existing reference framework. The underlying intention is to prepare a wider application of single components beyond the current online scheduling application. Examples would be a usage of the simulation model and the interface methods for the optimization of factory layouts or application in intralogistics as described in more detail in the last chapter 5.

In the following, the allocation and purposes of the software components in the different Internet of Production layers are described briefly. The system level provides the scheduling problem and the shopfloor data. The interface methods act as a middleware+ that is capable of connecting the different software components. The scheduling server and the data model reside at the integration layer to provide multi-modal access to the online scheduling agents. They are responsible for decision-making and autonomous actions to provide the scheduling assignments in the smart expert layer.

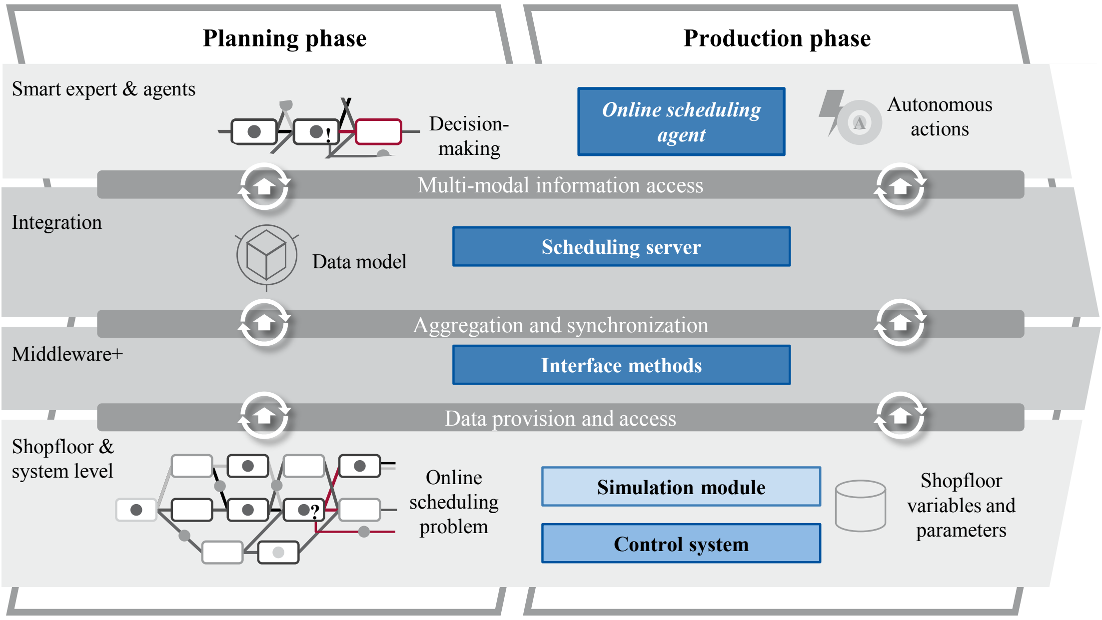


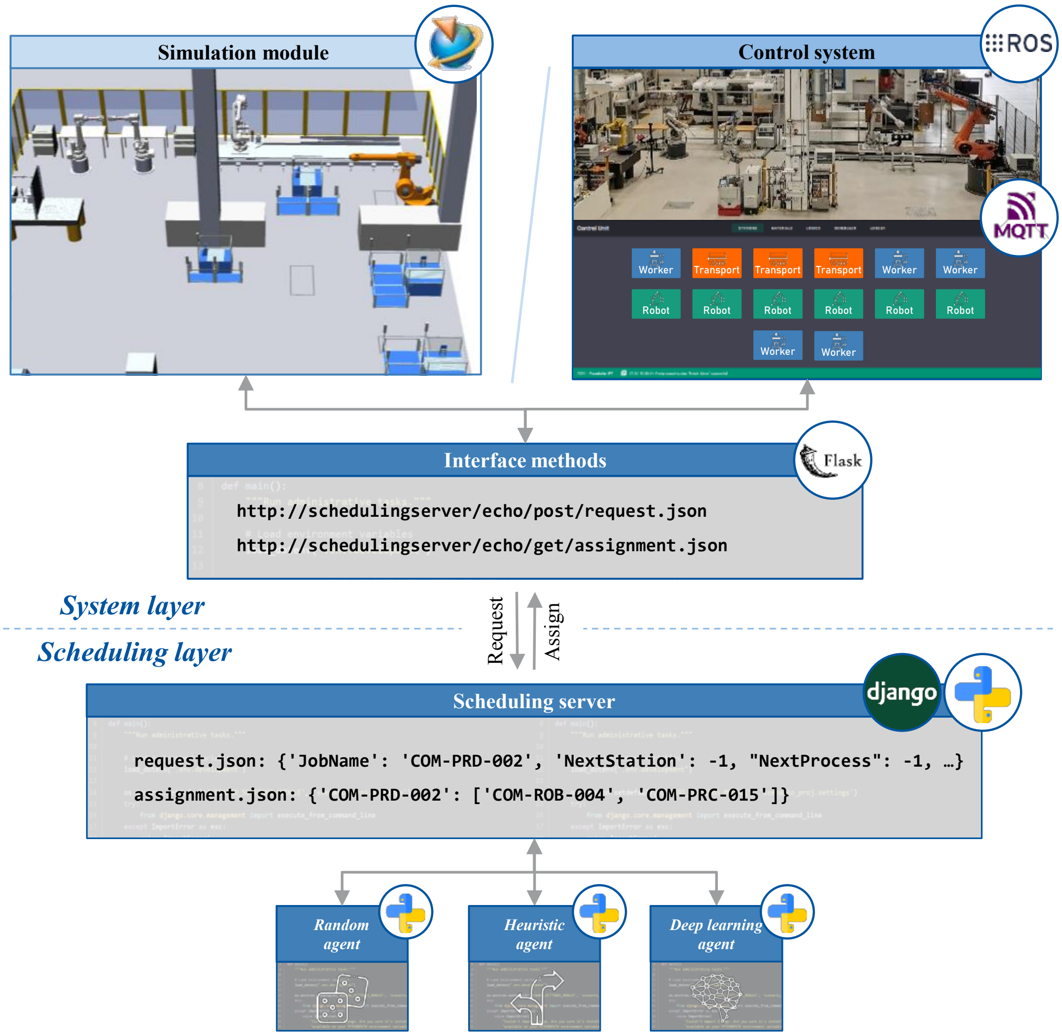
Figure 2: Incorporation of the modular software components in the reference framework Internet of Production for the production technology perspective (based on [35]).

4. Implementation and Functional Testing

The components and interface methods from the conceptual architecture presented in Figure 1 are implemented in infrastructure of the machine hall of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. The machine hall is equipped with several robots, manual work stations and automated guided vehicles. The software infrastructure consists of a Robot Operating System (ROS) middleware and the Message Queuing Telemetry Transport (MQTT) protocol for distributed communication. The control system COPE, developed by the Fraunhofer Institute for Production Technology (IPT) [27] is connected to the MQTT communication broker.

In Figure 3 in the top right corner, a section of the machine hall and the COPE user interface are visualized. The machine hall is represented in a discrete-event simulation model, created with Tecnomatix Plant Simulation, as shown on the top left.

Both, the simulation model and the COPE control system integrate the interface methods. The interface is set up with the Flask micro-framework, offering standardized RESTful Hypertext Transfer Protocol (HTTP) commands for transferring JavaScript Object Notation (JSON) files for requesting and assigning scheduling decisions with the HTTP methods POST and GET, respectively. The scheduling server was created with the high-level Python web framework Django, offering classes and methods for setup up the addressable web server. Within the scheduling server the functionalities of resolving deadlocks by enforcing sub-optimal scheduling decisions and solving trivial scheduling problems are also implemented in Python. In Figure 3 two excerpts of the content of request and assignment JSON files are shown. For the request, the file contains the name of the to-be-scheduled job as an identifier and the NextStation and NextProcess variables with the value -1, indicating a scheduling demand. Not listed in the excerpt is the current information about the assembly system with all current status of resources (robots, transport, manual work stations, etc.), other jobs and processes. In addition to the live date, static data that contains the overall system setup, such as locations of work stations, is included in the request. The assignment reply consists of the job name (COM-PRD-002) and the station-process combination (COM-ROB-004 and COM-PRC-015). The nomenclature of these strings follows a JSON scheme, i.e. three letters or words and predefined letters, e.g. for the robot station. This scheme is defined the meta model, to allow standardized data formatting (according to [33]).



```
http://schedulingserver/echo/post/request.json
http://schedulingserver/echo/get/assignment.json
```

```
request.json: {'JobName': 'COM-PRD-002', 'NextStation': -1, "NextProcess": -1, …}
assignment.json: {'COM-PRD-002': ['COM-ROB-004', 'COM-PRC-015']}
```

Figure 3: Overview of the implemented software components, following the modular software architecture.

The implemented scheduling agents can be chosen with a parameter in the scheduling request and the scheduling server is forwarding the request to the respective scheduling agent. The deep learning online scheduling agent is implemented according to the descriptions in Göppert et al. (2020, 2021). [36,37] The random agent chooses scheduling assignments randomly and the heuristic agent (also explained in [36]) minimizes the waiting and transportation time for the next station-process combination. The random and heuristic agents serve as reference agents to evaluate the scheduling performance of the deep learning agent. All scheduling agents were implemented in Python using open libraries (e.g., NumPy, scikit-learn, Keras).

The functionality of all components and interfaces was tested with qualitative validation measures, i.e. single error handling checks and large-scale simulation experiments. The qualitative validation results with corresponding inputs and functions are presented in Table 1.

Table 1: Qualitative validation of interfaces and scheduling server functionality.

Input description	**Function**	**Result**
Data model content not schema-compliant in request (e.g. 'COM-PRODUCT-…')	Scheduling server request consistency check	Scheduling server error message: "Data model incorrect"
Missing data model content in request (e.g. no NextProcess value given)	Scheduling server request consistency check	Scheduling server error message: "Data model incorrect"
Sending scheduling request with one possible station-process combination	Scheduling server functionality check	Assignment is consistent with the possible station-process combination, no deviation could be observed
Response message from scheduling agent incorrect	Scheduling server agent message check	Scheduling server error message: "Scheduling agent error"
Sending the same scheduling requests from a simulation model and control system	Interface function consistency check	Identical scheduling results for simulation and control system
Not existing algorithm name requested (e.g. "test_agent")	Scheduling server request consistency check	Scheduling server error message: "Algorithm incorrent"

In addition to the qualitative testing, a large number of test scenarios with the simulation model was executed as a large-scale validation. As shown in Göppert et al. (2020) 10,944 simulation scenarios, i.e. full simulation runs with 180 scheduling decision points per scenario were generated with the automated scenario analysis tool. [36] In total, the $1.97\text{x}10^6$ scheduling requests were processed by the scheduling server and the scheduling agents. The scheduling assignments were successfully retrieved by the simulation leading to validated scenario outputs in the form or reasonable performance indicators.

5. Conclusion and Outlook

In this last chapter, a conclusion on the research question with the key findings and an outlook on potential applications and research opportunities is provided. The research question – *How can the seamless connection of various online scheduling agents with planning and control systems be enabled?* – can be answered as follows: The standardized interfaces, the dedicated scheduling server and the overall validated modular software architecture enable the planning and control with various online scheduling agents.

From the presented conceptual architecture and implementation, the key findings and additions to the knowledge base are:

1. The dedicated scheduling server software module enables the encapsulation of scheduling decisions and externalization from simulation or control systems to, for instance, enable independent scheduling agent development.
2. Advantages of the scheduling server are the central deadlock resolution to avoid repeated implementation in the scheduling agents for trivial scheduling task solution and to reduce unnecessary communication.
3. Web-based interfaces allow for the system-agnostic (simulation or control) handling of scheduling requests and assignments and, therefore, enable switching seamlessly from simulating during planning to the control of a running system.

4. Applying data modelling standards such as schemes and meta-models facilitates the interoperable communication via data models and the modularization of components.
5. The mentioned open-source libraries for web-server development give researches and practitioners from production technology potential support in creating a comparable modular scheduling architecture.

Besides these benefits and learnings, also drawbacks of the presented architecture exist. A downside of the modularity are the additional communications efforts between the system layer, the scheduling server, and the scheduling agents. The time for the additional communication results in an overhead waiting time for each scheduling request. For simple scheduling requests, the communication overhead might reduce the performance, especially for discrete-event simulations that seek to simulate large numbers of scenarios quickly. Extra external communication efforts for each scheduling decision, can significantly increase the total simulation time. Also, the development and implementation costs of the dedicated scheduling server and the interface methods have to be considered. Proprietary software for simulation or control might also not allow the integration of the interface methods.

Further potential applications of the proposed software architecture and its components can be found outside assembly systems, wherever online scheduling is applied. The data model for resembling the real system must be adapted to the application, but the functionality of the components is independent of that. A field with high potential is intralogistics, due to the short transport times and ad-hoc request of transportation, which cannot be scheduled ahead in large and complex systems. Further exemplary fields outside the manufacturing domain are airplane terminal allocation, scheduling of patients to hospital resources and assigning data computing tasks to processing units.

Further research opportunities, besides the broader application in the above-mentioned fields, comprise the measurement of these external communication times depending on the various communication protocols or web server frameworks. With these results, the impact of external communication in contrast to the benefits of modular scheduling agents can be investigated. Also, the communication protocol that supplies the fastest response times can be chosen. In general, the long-term application of the modular scheduling architecture in real production with a control system could lead to more insights about potential interface improvements and could validate the postulated benefits. Furthermore, the long-term application would increase the architecture's maturity. Eventually, standardization committees could develop a specific reference architecture that follows domain-wide definitions of interfaces and data models.

Acknowledgements

We would like to thank the BMWi, DLR-PT and our partners for their kind support. This work is part of the research project "AIMFREE", which is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the "Directive on a joint funding initiative to support research and development in the field of electro mobility" (funding code: 01MV19002A) and supported by the project management agency German Aerospace Center (DLR- PT). The authors are responsible for the content. More information about the current research activities within the project AIMFREE at www.aimfree.wzl.rwth-aachen.de.

References

[1] Lanza, G., Nyhuis, P., Fisel, J., Jacob, A., Nielsen, L., Schmidt, M., Stricker, N., 2018. Wandlungsfähige, menschzentrierte Strukturen in Fabriken und Netzwerken der Industrie 4.0. acatech Studie) München," Herbert Utz Verlage.

[2] Rauch, E., 2013. Konzept eines wandlungsfähigen und modularen Produktionssystems für Franchising-Modelle. Fraunhofer-Verlag, Stuttgart.

[3] Belhadi A., Kamble, S. Jabbour, C., Gunasekaran, A., Ndubisi, N.O., Venkatesh, M., 2021. Manufacturing and service supply chain resilience to the COVID-19 outbreak: Lessons learned from the automobile and airline industries. Technological Forecasting and Social Change 163, 120447.

[4] Hees, A., 2016. System zur Produktionsplanung von rekonfugrierbaren Produktionssystemen. Dissertation, München.

[5] Nyström, M., Jouffray, J.-B., Norström, A.V., Crona, B., Jørgensen, P.S., Carpenter, SR, Bodin, Ö., Galaz, V., Folke, C., 2019. Anatomy and resilience of the global production ecosystem. nature 575 (7781), 98–108.

[6] Greschke, P.I., 2015. Matrix-Produktion als Konzept einer taktunabhängigen Fließfertigung. Dissertation, Braunschweig.

[7] Schönemann, M., Herrmann, C., Greschke, P., Thiede, S., 2015. Simulation of matrix-structured manufacturing systems. Journal of Manufacturing Systems 37, 104–112.

[8] Kern, W., Rusitschka, F., Bauernhansl, T., 2016. Planning of Workstations in a Modular Automotive Assembly System. Procedia CIRP 57, 327–332.

[9] Ramsauer, C., Rabitsch, C., 2015. Agile Produktion - Ein Produktionskonzept für gesteigerten Unternehmenserfolg in volatilen Zeiten, in: , Industrial Engineering und Management. Springer Fachmedien Wiesbaden, pp. 63–81.

[10] Hüttemann, G., Göppert, A., Lettmann, P., Schmitt, R.H., 2017. Dynamically Interconnected Assembly Systems, in: Schmitt, R., Schuh, G. (Eds.), 7. WGP-Jahreskongress Aachen, 5.-6. Oktober 2017. Apprimus Wissenschaftsverlag, Aachen, pp. 261–268.

[11] Heilala, J., 2022. Modeling and Simulation for Decision Making in Sustainable and Resilient Assembly System Selection. Scandinavian Simulation Society, 180–188.

[12] Wu, X., Zhao, J., Tong, Y., 2018. Big Data analysis and scheduling optimization system oriented assembly process for complex equipment. IEEE Access 6, 36479–36486.

[13] Pinedo, M.L., 2014. Scheduling: Theory, Algorithms, and Systems, 5th ed. Springer.

[14] Usuga Cadavid, J.P., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A., 2020. Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. Journal of Intelligent Manufacturing 31 (6), 1531–1558.

[15] Morariu, C., Morariu, O., Răileanu, S., Borangiu, T., 2020. Machine learning for predictive scheduling and resource allocation in large scale manufacturing systems. Computers in Industry 120, 103244.

[16] Chaudhry, I.A., Khan, A.A., 2016. A research survey: review of flexible job shop scheduling techniques. International Transactions in Operational Research 23 (3), 551–591.

[17] Fang, Y., Peng, C., Lou, P., Zhou, Z., Hu, J., Yan, J., 2019. Digital-twin-based job shop scheduling toward smart manufacturing. IEEE Transactions on Industrial Informatics 15 (12), 6425–6435.

[18] Minguillon, F.E., Stricker, N., 2020. Robust predictive-reactive scheduling and its effect on machine disturbance mitigation. CIRP Annals 69 (1), 401–404.

[19] Rinciog, A., Mieth, C., Scheikl, P.M., Meyer, A., 2020. Sheet-Metal Production Scheduling Using AlphaGo Zero. Conference On Production Systems and Logistics, 1–10.

[20] Dorri, A., Kanhere, S.S., Jurdak, R., 2018. Multi-Agent Systems: A Survey. IEEE Access 6, 28573–28593.

[21] Leusin, M.E., Kück, M., Frazzon, E.M., Maldonado, M.U., Freitag, M., 2018. Potential of a Multi-Agent System Approach for Production Control in Smart Factories. IFAC-PapersOnLine 51 (11), 1459–1464.

[22] Park, B., Jeong, J., 2020. A CPS-Based IIoT Architecture Using Level Diagnostics Model for Smart Factory, in: Computational Science and Its Applications - ICCSA 2020. Springer International Publishing, Cham, pp. 577–587.

[23] Helo, P., Phuong, D., Hao, Y., 2019. Cloud manufacturing – Scheduling as a service for sheet metal manufacturing. Computers & Operations Research 110, 208–219.

[24] Oluyisola, O.E., Bhalla, S., Sgarbossa, F., Strandhagen, J.O., 2022. Designing and developing smart production planning and control systems in the industry 4.0 era: a methodology and case study. J Intell Manuf 33 (1), 311–332.

[25] Jennings, N.R., 2000. On agent-based software engineering. Artificial Intelligence 117 (2), 277–296.

[26] Zhou, T., Tang, D., Zhu, H., Zhang, Z., 2021. Multi-agent reinforcement learning for online scheduling in smart factories. Robotics and Computer-Integrated Manufacturing 72, 102202.

[27] Jung, S., Ochs, J., Kulik, M., König, N., Schmitt, R.H., 2018. Highly modular and generic control software for adaptive cell processing on automated production platforms. Procedia CIRP 72, 1245–1250.

[28] Krockert, M., Matthes, M., Munkelt, T., 2021 - 2021. Agent-based Decentral Production Planning and Control: A New Approach for Multi-resource Scheduling, in: Proceedings of the 23rd International Conference on Enterprise Information Systems. 23rd International Conference on Enterprise Information Systems, Online Streaming, --- Select a Country ---. 26.04.2021 - 28.04.2021. SCITEPRESS - Science and Technology Publications, pp. 442–451.

[29] Vieira, M., Moniz, S., Gonçalves, B.S., Pinto-Varela, T., Barbosa-Póvoa, A.P., Neto, P., 2021. A two-level optimisation-simulation method for production planning and scheduling: the industrial case of a human–robot collaborative assembly line. International Journal of Production Research, 1–21.

[30] Aggarwal, S., 2019. Flask Framework Cookbook: Over 80 proven recipes and techniques for Python web development with Flask. Packt Publishing Ltd.

[31] Cerny, T., Donahoo, M.J., Trnka, M., 2018. Contextual understanding of microservice architecture: current and future directions. ACM SIGAPP Applied Computing Review 17 (4), 29–45.

[32] Rabe, M., Spieckermann, S., Wenzel, S., 2008. Verifikation und Validierung für die Simulation in Produktion und Logistik: Vorgehensmodelle und Techniken. Springer Science & Business Media.

[33] Göppert, A., Grahn, L., Rachner, J., Grunert, D., Hort, S., Schmitt, R.H., 2021. Pipeline for ontology-based modeling and automated deployment of digital twins for planning and control of manufacturing systems. Journal of Intelligent Manufacturing.

[34] Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51 (11), 1016–1022.

[35] Brecher, C., Klocke, F. Schmitt, R.H., Schuh, G. (Ed.), 2017. Internet of Production für agile Unternehmen: AWK Aachener Werkzeugmaschinen-Kolloquium. Apprimus Verlag.

[36] Göppert, A., Mohring, L., Schmitt, R.H., 2021. Predicting performance indicators with ANNs for AI-based online scheduling in dynamically interconnected assembly systems. Production Engineering (Journal) 15 (5), 619–633.

[37] Göppert, A., Rachner, J., Schmitt, R.H., 2020. Automated scenario analysis of reinforcement learning controlled line-less assembly systems. Procedia CIRP 93, 1091–1096.

Biographies

Amon Göppert (*1992) is head of the department of Model-based Systems at Laboratory of Machine Tools and Production Engineering of RWTH Aachen University (WZL) since 2021. From 2017 to 2021 he was a research associate at WZL.

Jonas Rachner (*1995) is a research associate at the department of Model-based Systems at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2019.

Lea Kaven (born Grahn) (*1995) is research associate at the department of Model-based Systems at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2020.

Prof. Dr.-Ing. Robert H. Schmitt (*1961) is head of the Chair of Production Metrology and Quality Management at Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University and board member of the Fraunhofer Institute for Production Technology (IPT).

3rd Conference on Production Systems and Logistics

Investigation Of Wire Mark Reading Methods To Support Automatic Quality Control

Stefanie Bartelt[1], Bernd Kuhlenkötter[1]
[1]*Ruhr University Bochum, Chair of Production Systems, Bochum, Germany*

Abstract

During the assembly of a control cabinet, a worker obstructs many individual configured wires. To distinguish these wires, a printer plots an identifying text on each end of the wires. However, due to the shape of the wires and the printing process, the quality of these markings is often too low, and it is hard or impossible to read the marking. Common reasons are a low contrast or a blurred text. By now, there is no quality check of the marking after a crimping machine produced the wire. This paper investigates methods for wire mark reading that is required to estimate the quality of the marking. By using optical character recognition, the likeliness that a worker can read the marking must be computed. In the final solution, the quality check of the marking will be implemented within an automated quality check that is located after the printing process. With this, the crimping machine can then discard wires of low quality and reproduce them instantly.

Keywords

Quality control; preconfigured wires; wire marking; machine learning; optical character recognition.

1. Introduction

Control cabinets are quite common in different domains. Although the specific cabinets can differ largely, their general setup is very similar. Each control cabinet contains a mounting plate in the back. On this plate, wire ducts and top hat rails are fixed. The top hat rails simplify the assembly of further components that are required by the customer. Nevertheless, components may also be fixed directly to the mounting plate. To connect the deployed components, wires with a different color and a different cross section are utilized. The cross section is selected according to the maximum current and the color is often defined by the function of the connection, e.g., signal, power line, or ground. [1]

Figure 1 depicts an image of a fully assembled control cabinet. The top hat rails are filled with components and cannot be seen. The wires are guided from the components' connections in a preferably short way into a nearby wire duct. Considering an average control cabinet, the cabinet embeds about a hundred or more wires. As the figure indicates, the wire ducts often contain a larger number of different wires. Hence, the identification of a certain wire can be quite challenging.

To simplify this identification, markings can be printed on the wires. A common practice is to use the source as well as the target. Each component has a reference designator that is unique within a project. Standards, e.g., the EN 81346, specify rules for the naming and are commonly used in industry. However, obsolete specifications like the DIN 40719 are still used for naming. For instance, a terminal block may be identified by the string "=0815+LA-X10:2-". Thereby, the first part starting with the equal sign indicates the facility,

DOI: https://doi.org/10.15488/12175

and the second part starting with a plus sign indicates the location. Both do not change for a control cabinet and can be omitted within a marking. Hence "-X10:2-" would be a meaningful marking for one end of a wire and identifies a certain component in the setup. Let us assume that "-X10:SPE" is another marking for the other end of the wire. As soon as a worker reads the end marking, he will know where to connect the wire to. However, the length of the wire might be up to a few meters and to identify a wire at an arbitrary position, intermediate markings are added to the wire. Angle brackets gives a hint to the direction. In the example above, "-X10:2 < > -X10:SPE" is such an intermediate marking. Following the wire to the left, one will end up at the end with the marking "-X10:2", and so on. Of course, other rules are possible, but creating markings according to these rules is simple and effective.

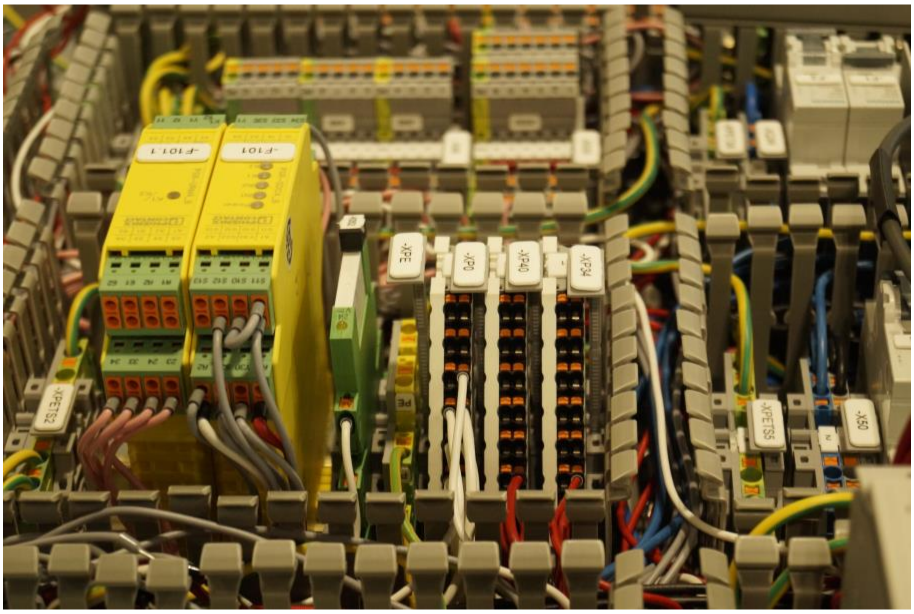


Figure 1: Partial image of a fully assembled control cabinet. Only the covers of the wire ducts are missing.

2. Problem statement

During assembly as well as for maintenance, it is important to identify a certain wire. For this, a sufficient quality of the above-mentioned marking is important. However, an inkjet printer applies the marking to the wire. This process is very susceptible to errors, and a proper quality check is missing. The most common errors are a blurred or faded print. A suitable solution for a quality check would be the utilization of optical character recognition (OCR). The OCR algorithm can identify the marking's text, which is then compared to the known one that is applied by the printer. To illustrate the problem, figure 2 depicts some wires with different diameter and color. As shown, also the color of the ink may be different. Due to this, the contrast of the printing and the color of the wire can be very low.

Although such an inspection can be done in a defined environment, some aspects must be considered for the quality check:

- The wires are most likely not in a straight line, but they are usually slightly bent. As mentioned above, there is a defined environment for the inspection. This can prevent greater bending, but some remains. With respect to the character recognition, this will be a minor issue.
- The printer applies the marking only from one direction. Hence, due to the shape of the wires, the marking may not be fully visible when using only one camera for inspection. Due to a twist of the wire, it might even happen that some characters are fully visible, and some are not.

- The font of the printer is a dot-matrix font. While common serif or sans-serif fonts and even handwritten text is state-of-the-art, dot-matrix fonts are very hard to identify. The main reason for this failure seems to be the disconnected points that prevent common methods to work properly.

Figure 2: Image of wires with different size and colors.

3. Related work

Optical character recognition has been subject of research for many decades. The general approach is to acquire an image, detect text lines and positions of single characters, and identify the characters found [2]. Finally, the identified characters are grouped to words. In addition, the image may be pre-processed, and the identified text can be post-processed to reduce errors. However, for the above-described problem, only the steps up to the building of words is relevant, because to decide whether the quality is sufficient or not, this identified text must be compared with a known one. Furthermore, the building of words is very simple since there is only one text line and white spaces can be neglected.

In recent years, the utilization of machine learning for OCR applications has increased. Although the main steps are the same as described above. Thereby, different models for text detection can be selected, i.e., an object detection model or a text instance segmentation model. Eskenazi et al. gives a review on several segmentation algorithms [3]. A succeeding transcription model yields then to the final text. Furthermore, a character instance segmentation model can be used for both the text detection and text transcription [4]. There are two major applications for OCR. The first one is the text recognition of printed documents, for instance, books or invoices. Such documents contain a huge number of characters and have a good structure, i.e., lines and columns of text. The second application is text recognition in real life. Thereby, it is most likely that texts are rotated or distorted, and a main issue is to find the positions of the characters within an image [5].

In [6], Zheng et al. present a smart assistance system based on OCR. The system combines augmented reality methods and visual inspection methods. With this, the system can identify certain wires and present corresponding information to the user. Although the system can detect text printed to the wires, the text font is a sans-serif font and off-the-shelve methods are able to recognize them.

Dot-matrix characters are formed from single dots that are not connected. In addition, the appearance of a character may change significantly already if a single dot is missing. As a result, recognition of dot-matrix fonts is different from other fonts [7]. Approaches are either an enhanced pre-processing, like connecting the dots, or sophisticated training-based methods. An efficient solution to find dots within an image is the computation of the cross correlation of the image with a given image of a single dot. For this, fast algorithms exist, e.g. [8]. In [9], the authors proposes a combination of pre-processing to identify regions of interest and a convolutional neural network for character recognition. Nevertheless, there is no general solution for recognition of dot-matrix fonts.

Szajna et al. are developing a system to read wire markings by means of artificial intelligence, i.e., a deep neural network [10]. The presented solution takes a picture from a wire inserted into the system. With this, advanced methods recognize the wire marking. Thereby, the focus is on identifying any character including faulty ones. Although the system may be adapted to a quality control, the project does not analyze different fonts used for the marking, since it is assumed that the certain specification for the labeling varies from company to company. Even though the examined wires were marked with a dot-matrix font, the individual dots are sufficiently large and connected, which makes recognition much easier.

4. Wire mark reading

There are two main steps to read a wire marking. The first step is the image acquisition with an optional pre-processing of the image. The second step is the character recognition. From the related work, a setup with a fixture for the wire, illumination, and a camera that takes images with a medium resolution is proven to yield good results. To find an appropriate solution, different variants of pre-processing as well as various character recognition solutions were compared.

4.1 Acquisition setup

The first step in identifying the marking was to build a well-defined environment. For this, a camera module and illumination are located above a frosted glass plate. Although the frosted glass creates a slight reflection of the camera module, it eliminates almost completely shadows of the wire. Opaque plates are mounted to the sides to eliminate effects from outside. The wires can be installed through holes on two opposite sides. The OV2640 camera module is controlled by an ESP32 microcontroller. Next to the camera module, a white color LED is located that is used for illumination. Via a serial connection, a dedicated application can acquire an image with a resolution of 1024x768 pixel. Figure 3 shows the development state of the hardware and software prototype.

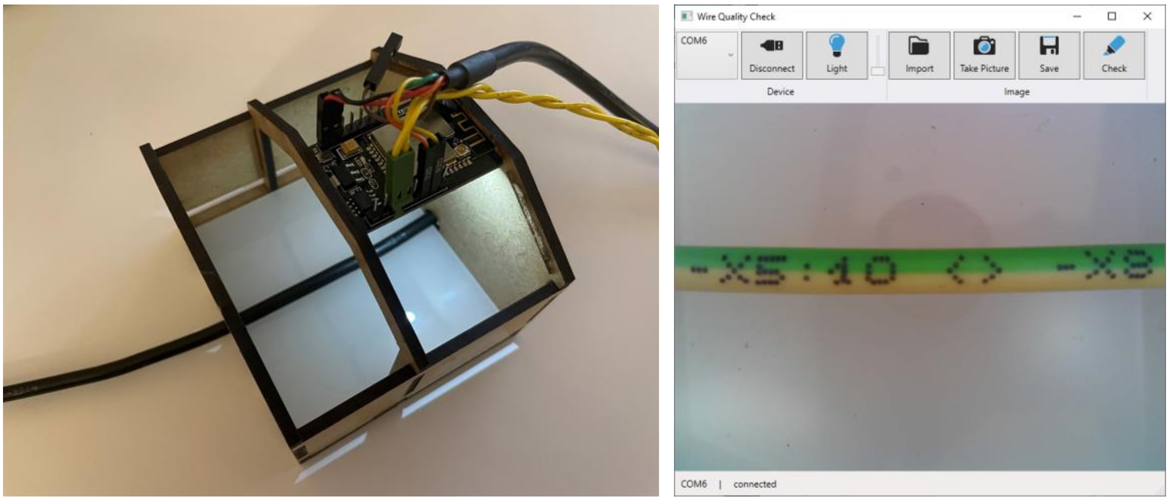
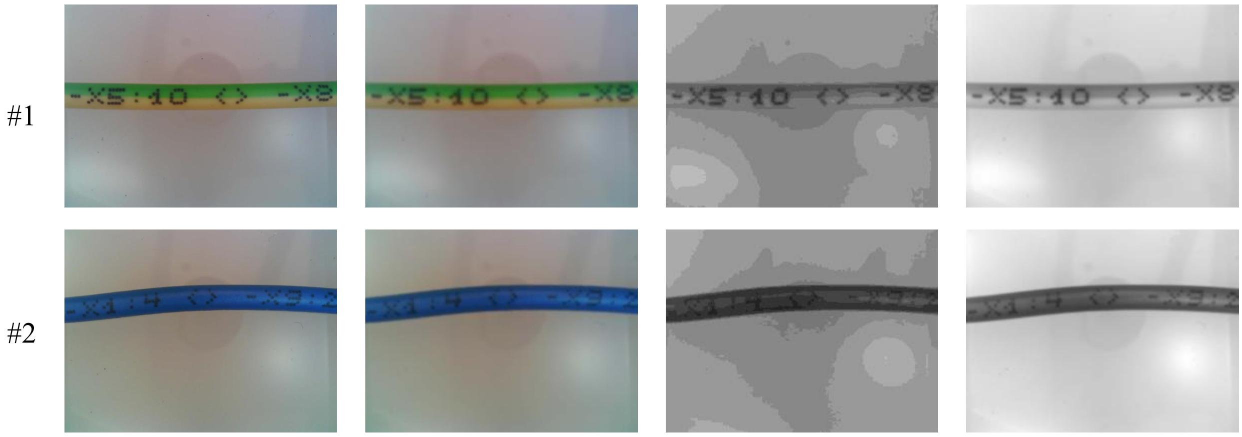


Figure 3: Photo of the hardware (left) and software (right) prototype to identify the markings.

4.2 Pre-processing

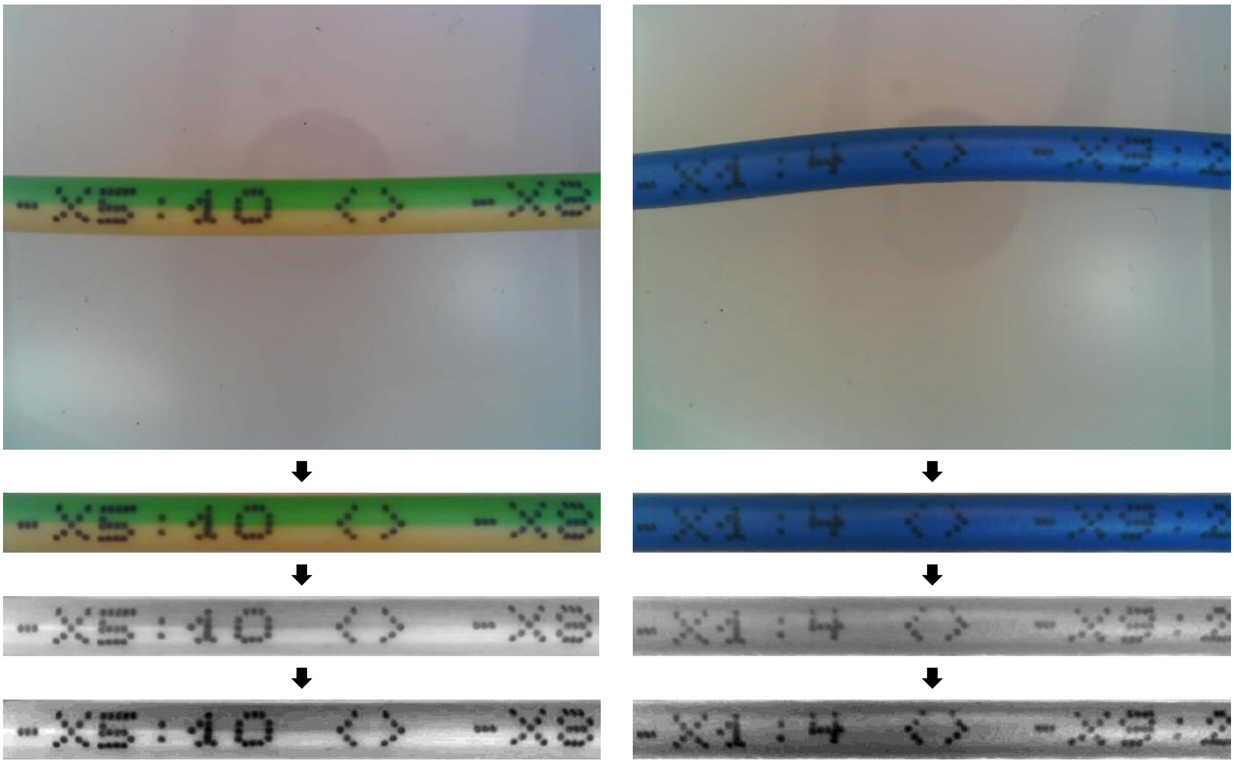
With the above-described setup, 65 images of different wires were taken and processed. In a first test, it turns out that gray scaling, blurring, and cross correlation increases the recognition. Figure 4 gives two examples of an original image as well as three pre-processed variants. The first variant applies a scatter filter before a Gaussian blur. The next one applies the same filter and adds a gray scaling. The last variant is a cross correlation with a black dot on a white background. Other variants were also tested, but they do not yield better results.



(a) (b) (c) (d)

Figure 4: Two examples of an original image (a), blurred image (b), blurred grey-scaled image (c), and cross-correlation image (d).

As described in the next section, this pre-processing improves the character recognition. However, the effect is limited. Furthermore, recognition for wires with a low contrast, for instance a blue wire with black font such as example #2 in figure 4 shows, was not possible in any case.



(a) (b)

Figure 5: Three applied steps of pre-processing consisting of straitening, color removal, and normalization.

To improve the contrast, the three-step pre-processing was selected, as figure 5 depicts. The first step is to straighten the wire. For this, the contour of the wire is detected by applying a Canny edge detection. With this, a perspective transformation computes a straight image of the wire. The second step estimates the wire color and removes it by using formula (1). $\overline{\mathrm{Red}_y}$ is the normalized mean value of line y and may has a value from zero to one. $\mathrm{Gray}_{x,y}$ is the gray value of the pixel at line y and column x. The other colors are equivalent. The last step is a linear normalization to increase the contrast.

$$Gray_{x,y} = \left(1 - \overline{\mathrm{Red}_y}\right) \cdot \mathrm{Red}_{x,y} + \left(1 - \overline{\mathrm{Green}_y}\right) \cdot \mathrm{Green}_{x,y} + \left(1 - \overline{\mathrm{Blue}_y}\right) \cdot \mathrm{Blue}_{x,y} \qquad (1)$$

4.3 Character recognition

The pre-processed images as well as the original image were tested with several OCR engines. Table 1 states the result of the recognized text for the given example. Further engines were also tested, e.g., Tesseract, IronOcr, and Aspose OCR. Except of Matrox SureDotOCR, these engines do not compute viable results for any of the image variants. As shown, only one engine can recognize text in the original image. The detected text corresponds to the “X8” that was recognized as a 180° rotated “SX”. Also, the blurred image has a bad performance. The grey scaled variant has a quite good recognition compared to the colored one, and even better than the cross-correlation image. Interestingly, the characters “<” and “>” were not detected by any engine. Nevertheless, all tested OCR engines have failure rates that are far from acceptable. An exception is SureDotOCR. This engine is specialized for reading dot matrix fonts and can detect the marking properly. For this, the used 5x5 dot font is defined within the engine. A major drawback is the requirement to specify the expected number of characters. As soon as this number does not match to the image, the results are rather bad. For example, when trying to read 11 characters in the example given in figure 5 a, the correct string of “-X5:10 < > -X8” was read. When trying to read 12 characters, the engines gives “--:P-P-:--:-”.

Table 1: Character recognition results of different engines.

image	Google vision	OCRSpace	fintract	Matrox SureDotOCR
Figure 4 #1a	SX	*no text detected*	*no text detected*	-X5:10 < > -X8
Figure 4 #1b	*no text detected*	-X5 10	*no text detected*	-X5:10 < > -X8
Figure 4 #1c	- X5:10 -XB	-X5:1O 8	- X5:10 -XB Xe	-X5:10 < > -X8
Figure 4 #1d	SS: Aus-XS	-X5 1O	- XE: 40 -XE	-X5:10 < > -X8
Figure 4 #2a	*no text detected*	*no text detected*	*no text detected*	*no text detected*
Figure 4 #2b	*no text detected*	*no text detected*	*no text detected*	*no text detected*
Figure 4 #2c	*no text detected*	*no text detected*	*no text detected*	*no text detected*
Figure 4 #2d	*no text detected*	*no text detected*	*no text detected*	*no text detected*
Figure 5 a	s	*no text detected*	000	-X5:10 < > -X8
Figure 5 b	တစ်မျို	*no text detected*	*no text detected*	-X1:4 < > - X9:

5. Conclusion and future work

Within this paper, the problem of recognizing text on electrical wires was stated. An optical character recognition algorithm should identify the printed characters, and the resulting text can be compared with the known marking text. If both texts match, the quality of the printing is sufficient. While most commonly available engines cannot read the marking properly, one engine that is dedicated to dot-matrix fonts yield appropriate results. Although this library is suitable for a quality check of a known marking, a general detection is not possible due to limitations of the library in terms of flexibility of the number of characters. Hence, further developments are required, which will be done in future work. Additionally, further work will elaborate the challenge that all sides of the wire must be considered. By now, the wires are manually rotated to ensure the marking to be on top.

Acknowledgements

This Project is supported by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag (funding number: KK5055201MS0).

References

[1] Tempel, P., Eger, F., Lechler, A., Verl, A., 2017. Schaltschrankbau 4.0: Eine Studie über die Automatisierungs- und Digitalisierungspotentiale in der Fertigung von Schaltschränken und Schaltanlagen im klassischen Maschinen- und Anlagenbau. Universität Stuttgart. Institut für Steuerungstechnik der Werkzeugmaschinen und Fertigungseinrichtungen., Stuttgart.

[2] Islam, N., Islam, Z., Noor, N., 2017. A Survey on Optical Character Recognition System. http://arxiv.org/pdf/1710.05703v1.

[3] Eskenazi, S., Gomez-Krämer, P., Ogier, J.-M., 2017. A comprehensive survey of mostly textual document segmentation algorithms since 2008. Pattern Recognition 64, 1–14.

[4] Subramani, N., Matton, A., Greaves, M., Lam, A., 2020. A Survey of Deep Learning Approaches for OCR and Document Understanding, 15 pp. http://arxiv.org/pdf/2011.13534v2.

[5] Zhu, Y., Yao, C., Bai, X., 2016. Scene text detection and recognition: recent advances and future trends. Front. Comput. Sci. 10 (1), 19–36.

[6] Zheng, L., Liu, X., An, Z., Li, S., Zhang, R., 2020. A smart assistance system for cable assembly by combining wearable augmented reality with portable visual inspection. Virtual Reality & Intelligent Hardware 2 (1), 12–27.

[7] Endo, K., Ohyama, W., Wakabayashi, T., Kimura, F., 2015. Performance Improvement of Dot-Matrix Character Recognition by Variation Model Based Learning, in: Jawahar, C.V., Shan, S. (Eds.), Computer Vision - ACCV 2014 Workshops, vol. 9009. Springer International Publishing, Cham, pp. 147–156.

[8] Briechle, K., Hanebeck, U.D., 2001. Template matching using fast normalized cross correlation, in: Optical Pattern Recognition XII. Aerospace/Defense Sensing, Simulation, and Controls, Orlando, FL. Monday 16 April 2001. SPIE, pp. 95–102.

[9] Muresan, M.P., Szabo, P.A., Nedevschi, S., 92019. Dot Matrix OCR for Bottle Validity Inspection, in: 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP). 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania. 05.09.2019 - 07.09.2019. IEEE, pp. 395–401.

[10] Szajna, A., Kostrzewski, M., Ciebiera, K., Stryjski, R., Woźniak, W., 2021. Application of the Deep CNN-Based Method in Industrial System for Wire Marking Identification. Energies 14 (12), 3659.

Biography

Stefanie Bartelt (*1991) is a research assistant at the Chair of Production Systems at Ruhr University Bochum. She conducts research in the field of industrial robotics and is currently working on a project for the automated handling and quality control of bendable components for the wiring of control cabinets.

Bernd Kuhlenkötter (*1971) is Professor and head of the Chair of Production Systems (LPS) at Ruhr University Bochum. His research focuses on the planning, simulation and implementation of production systems and smart product-service systems. Prof. Kuhlenkötter is, among others, a member of the Scientific Society for Production Engineering (WGP) and the Scientific Society for Assembly, Handling and Industrial Robotics (MHI) as well as scientific director of the Research Center for the Engineering of Smart Product-Service Systems (ZESS).

3rd Conference on Production Systems and Logistics

Conceptualization Of The Use Of Artificial Intelligence For Interdependencies Analysis In Requirements Engineering*

Peter Burggräf[1,2], Georg Bergweiler[1], Falko Fiedler[1], Viktor Slawik[1], Jonas Recker[1]
[1]*Chair of Production Engineering / WZL of RWTH Aachen University, Aachen, Germany*
[2]*Chair of International Production Engineering and Management / Siegen University, Siegen, Germany*

Abstract

The efficiency in product development is largely determined by the quality of the requirements and the ability of the product design and production planner to analyze them. Interdependencies between multiple requirements identified at an early stage enable a sustainable design of the product as well as the corresponding production system by increasing process efficiency as well as the effectiveness of development processes. However, the necessary analysis of complex interdependencies between requirements of a product and the corresponding production system is time-consuming, error-prone, and highly inefficient when performed manually. Current development processes are based on such manual processes for analyzing requirements in natural language and must therefore be adapted.

This paper describes a methodical approach based on a semi-systematic literature review making the complexity of the interdependencies manageable by using existing approaches and methods in the field of model-based systems engineering (MBSE) as well as natural language processing (NLP). Thereby, a transition from informal requirements represented in natural language to analyzable and structured information, which enable interdependencies modeling for requirement chains, is described. A corresponding framework for analyzing interdependencies in the requirements engineering process is derived.

Keywords

Requirements Engineering; Artificial Intelligence; Natural Language Processing; Requirement Chains

1. Motivation

The increasing digitalization of the value chain and its technologies support a reduced time to market but also put the automotive industry under increasing pressure to innovate and adapt processes [1]. Increasing volumes of data for requirements engineering during planning processes call for automated approaches for structuring [2]. A large amount of product, as well as production-related data, can be assigned to planning and development processes [3]. Therewith, requirements engineering is a central part of the early design phase in the product development process (PDP) and bears the largest share of cost responsibility, accounting for two-thirds of the total[3][4]. Automated approaches are used increasingly to make requirements and change processes more efficient. In this context, the use of classification techniques and NLP enables a better understanding of relationships between requirements [2]. Technologies from the field of language and image understanding support planning and decision processes as well as the evaluation of large amounts of data [5]. Thereby, NLP enables the interaction between human language and digital information systems [6]. Methods from the field of Machine Learning (ML) have the ability, similar to the human brain, to identify patterns in large amounts of data and to react as a result [7]. While the human brain is characterized by

DOI: https://doi.org/10.15488/12143

---

* *The paper successfully passed the peer review process but was not presented at a conference.*

associative reasoning, the strength of algorithms lies in their distinct combinatory [7]. To better identify interactions between data and their properties, the use of models is suitable [2]. Models are used as bridges between undesired initial states and the corresponding desired final states and thus serve to solve problems [8]. The problem in requirements engineering emanates not only from the number of requirements but also from the multitude of sources for requirements. This is reflected in the form of information loss or ambiguity in requirements formulated in natural language [9]. The challenges of working in such an environment with a high level of process efficiency can be illustrated by a striking example as follows. The specifications of an aircraft engine contain requirements in the four-digit range [10]. However, a human being can only read an average of 170 words per minute [11]. If a requirement consists of fifteen words, each reading of the requirement specification takes at least one and a half hours. An analysis for correlations or defects such as contradictions etc. is not even included in the estimation. This example shows the enormous potential of automating parts of requirements engineering. It becomes obvious that there is an immense potential to reduce time and costs. During the requirements engineering the considered scope is increased from a part to larger assemblies and system-crossing (e.g. production) requirements are taken into account in addition to product requirements. Following not only inefficiencies but also the risk of quality losses due to a lack of overview of a large number of requirements and their interdependencies need to be avoided.

2. Problem statement and research task

As described, the manual and human management of requirements in product development is associated with a high expenditure of time. Lack of documentation and insufficient consistency of change status updates lead to improvisation in the requirements and change management process to a high degree. As a result, product changes take up more than one-third of the total resources in the design phase [12]. By applying data-based methods, parts of the requirements engineering process can be performed more efficiently. Although the necessary data basis already exists, the lack of formalization of technical requirements significantly limits modeling possibilities [2]. The demand for data-based methods is to combine the advantages of NLP, modeling, and ML. On this basis, increasing sets of requirements can be structured and, based on this, interactions can be identified and explicitly represented. The goal of this paper is to provide a framework for exploratory analysis of requirement chains, by applying methods of AI. In doing so, the focus is additionally on answering the following questions:

How can a process modeling framework for the transformation of separate requirements into transparent requirement chains be designed?

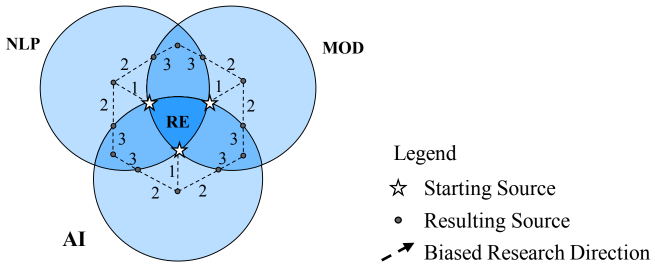
How can existing methods and approaches be classified along with the process modeling?

3. Methodical Approach

The content of this paper is part of a research project that uses the research methodology of Design Science Research (DSR). Following this approach, the development of an artifact takes place in the Design cycle between the Relevance cycle and the Rigor cycle [13,14]. Using an abductive approach through the use of an existing knowledge base in the form of existing literature, an artifact is generated through the described framework. This serves to situate the scientific theory within the research project of requirements chain generation.

Along with the process model of DSR, the activities of identifying the problem, defining requirements for the solution, and developing the artifact are thus addressed. A demonstration and evaluation take place in the context of the Rigor cycle since no application takes place in the business environment [15].

To find methods and approaches that fit the problem, a systematic approach for literature review is developed. The approach is based on the snowball principle. It uses the linkage of existing literature via source citations and acts as a cumulative search procedure based on this [16]. Intending to integrate the three identified foci equally, the snowball principle is extended. Figure 1 schematically depicts the developed methodology. The topics NLP, artificial intelligence (AI), and modeling have diverse overlaps in terms of content. Therefore three snowballs are shown, which form intersections in the center. The goal of the search is to link the three topics concerning the application field of requirements engineering, which is located in the center. To consider diverse approaches, the topic of the results is deliberately steered away from the core, and migrations between the fields are provoked. The result of the method is a literature collection with 48 results, whose thematic classification forms a symbolic hexagon around the core topic requirements engineering. From the identified results, the higher-level methods are extracted and promising specific approaches are analyzed.



Step	Branches Per New Result	New Results
0		3
1	3	9
2	2	18
3	1	18
Sum of Results		**48**

Figure 1: Systematic approach for the literature review

During the literature review, the STARLITE method is used to identify the most promising approaches. Only English titles with the publication year of 2016 or later are considered to focus on recent approaches [17].

The Web of Science database is used to identify a starting dataset. To focus on the core of the problem, a topic-specific search string is developed. This combines the identified topics using the logical operators AND as well as OR. The search string used is given below: TS=(("Requirements Engineering" OR "Requirement Management" OR "Requirements") AND ("MBSE" OR "SysML" OR "UML" OR "Relations") AND ("Artificial Intelligence" OR "Cased Based Reasoning" OR "Machine Learning" OR "Neural Networks") AND ("Natural Language Processing" OR "Natural Language Understanding" OR "Semantic Analysis" OR "Vectorization")). Three results that address the identified three main action areas in addition to requirements engineering are selected as starting sources (step 0 in fig. 1). Thereby, each of the results has the respective focus in one of the three fields. During the search for further results, the database Semantic Scholar is accessed. Starting from the initial dataset, the developed systematic snowball method is applied. Each identified source is considered as a starting point for the identification of further results.

4. The State of Research

The following section creates the foundation for conceptualizing the framework for the subsequent explanation and assignment of promising methods based on the systematically compiled literature. After describing the basics of requirements engineering, the fundamentals of modeling are discussed. Finally, basic concepts of AI are explained with a focus on natural language understanding.

4.1 Requirements Engineering

Requirements management is an essential part of the PDP. It represents the basis for product planning and development [3]. Requirements are functions or services that products must have to fulfill formal regulations such as standards or contracts [18]. They are defined at the beginning of the project and form a benchmark for later work in the product planning process. For this reason, requirements must be continuously checked and adapted if necessary [3]. To ensure the completeness and structure of the requirements collection, careful identification of all stakeholders involved is essential. Suppliers, laws and standards, production, sales, and controlling are sources of requirements. The most important sources are the market and the customer [19].

While documenting information, an appropriate format must be used in addition to appropriate labeling. Furthermore, a review of the input for suitability should take place [20]. In this context, the natural language documentation of requirements comes into focus. As a basis for documentation, the required performance of sourced products is recorded in a requirements specification. Based on the requirement specification, the contractor creates a requirement specification, which precisely defines the realization project to be developed [8]. Furthermore, documentation in the form of a requirements list is recommended, because requirements can be compared and prioritized [21]. High quality of the requirements documentation can be achieved by easily applicable formulation rules. These concern, among other things, sentence structure, sentence scope, and unambiguity of word choice [22]. To minimize the effort of the documentation process, requirements templates can be used. These provide a clear sentence structure for different types of requirements [9]. Attention must be paid to the initially identified requirements throughout the project. In addition to documentation, they also need to be communicated, maintained, and taken into account when evaluating concepts [19]. Following prioritization, they are compared over the entire development process and specifically introduced into the functional, activity, and construction levels based on the Munich Product Concretization Model [19]. Meanwhile, many external and internal factors have an impact on product development. For this reason, supporting the PDP with information technology is evident. This helps to ensure consistency of documentation, rapid exchange of results, and better traceability of activities [8].

4.2 Modelling

Models are an important part of engineering. They contain the foundations for databases in the form of logic, machine theory, and schemes [23]. Thereby, models are representations of a natural or artificial original based on abstractions [24]. During modeling, a limited set of attributes is transferred. Some attributes of the original are excluded. In return, new attributes are included in the developed model [24]. Models can be used as a basis for the development of products and support the solution of complex problems. Model-based development is based on models consisting of machine code and replaces handwritten texts [25]. Due to technological progress, products become more complex. This circumstance requires a more extensive system landscape [26]. Systems engineering (SE) supports the structuring of complex systems. It refers to the documentation of requirements concerning the holistic development picture [27]. MBSE combines the model character with SE. The goal is the transformation of heterogeneous product models into interconnected as well as consistent images of the products [28]. In addition to the model architecture and the behavior of components, requirements are also introduced at each abstraction level of a model [29]. In the environment of the MBSE, one speaks of a model as soon as it fits a given formal form. This is achieved as soon as structures and relationships can be derived automatically from given models [29]. For the construction of models in different industries, the modeling language UML was developed. Through the integrated extension mechanism, application-specific add-ons can be integrated [27]. In this context, SysML has emerged as a dialect of UML. SysML helps in describing structures, behaviors, and requirements of a system. The modeling language extends the repertoire of UML diagrams to be used by integrating requirements diagrams and associated relationship capabilities [29]. Diagrams visualize specific characteristics of the comprehensive model. The focus during filtering is set on defined viewpoints [25].

4.3 Artificial Intelligence

Nowadays, both requirements engineering and model-based development involve many manual steps. In the process, humans fall back on vague and incomplete information from their memory [5]. The growing amount of data due to more complex products further complicates human work. The application of specific knowledge is essential for efficient processes. Knowledge is created by interpretation from information, which is aggregated from data by working out relationships [7]. Therefore, accurate analysis of the data sets is essential. For this purpose, the enormous computing power of information systems is increasingly used for data processing [7]. The area of text processing is covered by NLP. By using algorithms, the transformation of natural language texts into machine-readable code is possible [30]. For this purpose, the linguistic levels of natural language texts must be analyzed [31]. The richness of semantics increases due to the integration of relations and relationships between individual words, thus representing the core of language understanding [6]. In addition to NLP, ML forms an important application area of AI research. Its efficient operation is based on a large amount of data [7]. For this reason, ML is closely related to the field of data mining. Data mining describes the extraction of knowledge from aggregated data. In the application field of language, data mining is also referred to as text mining. Text mining includes tasks such as classification of texts and identification of similar texts [32]. Different types of neural networks are used for fast and efficient information processing in the field of ML. They are characterized by their decentralized as well as the parallel structure and their learning ability [7]. The approach of neural networks is optimized by different types of learning initially or continuously. Supervised, unsupervised, partially supervised or self-supervised learning methods are used as well as reinforcement learning [6].

5. Conceptional Design of the Framework

The basics compiled above form the foundation for the systematic development of a solution space for the generation of requirement chains. With the help of higher-level concepts of model theory, a concept is first developed that describes the target states of the solution process. Subsequently, the concept is detailed by integrating specific solution increments and an application-oriented framework is presented.

5.1 Concept

The goal of the framework is to efficiently extract the existing relationships between separate technical requirements formulated in natural language from the diversity of a collection of requirements. Figure 2 shows the step-by-step procedure during the concept development of the framework.

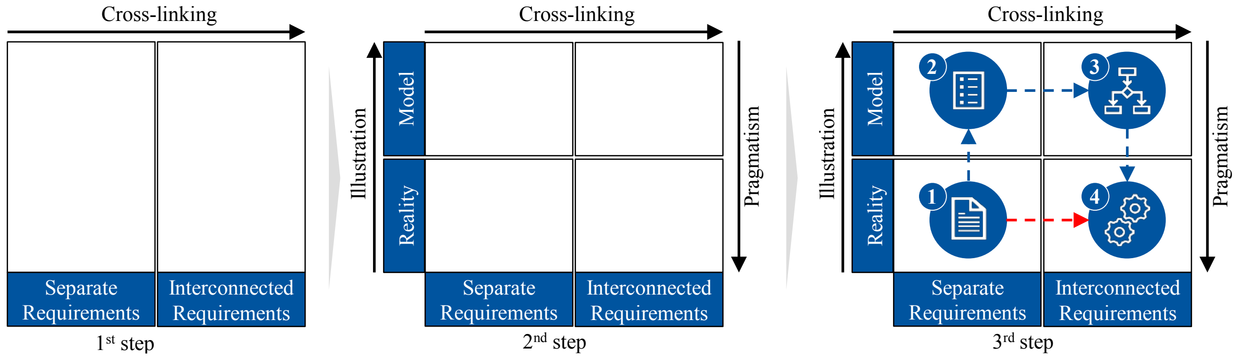


Figure 2: Step-by-step development of a concept for the regulatory framework

In a first step, a clear distinction between separate and non-structured requirements must take place. Before the existing relationships can be recognized automatically, each instance involved must fulfill the prerequisites required for this. The mechanism for linking the separate requirements is defined as cross-linking. In the second step, the model world is detached from reality. In the sense of the modeling idea, an issue can be mapped in the model level to open up new solution possibilities by abstracting a complex issue [24]. Thereby, illustration enables the transition from reality to the model [33,34]. In contrast, pragmatism serves to render the model in reality in a way that is understandable to the interpreter [24]. The two identified boundaries divide the solution space into four quadrants. Each quadrant is characterized by the unique combination of two characteristics determined in the first two steps. For each of the quadrants, a target state is defined in a third step. The overall goal is to analyze relationships between requirements to generate requirement chains. As a result, the separate natural language written requirements form the initial state (1) and the human-comprehensible representation of requirement chains (4) form the target state of the process. The goal of the transition into the model world is the structuring of the separate requirements (2). The goal of the networking of the separate requirements are cross-system requirement chains (3).

5.2 Detailed Framework

To detail the concept, the target states must be described by applicable processes. The transition between the first two target states is handled by modeling [33,34]. This comprises the transformation from continuous text, the structuring, and categorization as well as the formalized representation of the requirements formulated in natural language. The result is syntactically decomposed requirements in a machine-readable format. These contain additional information about the sentence-internal relationships between individual language elements. The subsequent cross-linking is the mechanism of the process evaluation. During this process, the identification of direct as well as indirect relationships between requirements takes place. Similarities between language elements are analyzed. In addition, higher-level relationships among subsystems are captured. The result are complex requirement networks, which contain a multiplicity of requirement chains. The discussion of relations takes place with the help of semantic information from semantic memory. This contains application-specific language relationships with increasing semantic richness and can be continuously extended. With the help of the process transfer the last target state can be reached [33,34]. By filtering information using perspectives and views as well as the automatic creation of diagrams, a human-readable visualization of the relationships is created. The resulting framework is given in Figure 3.

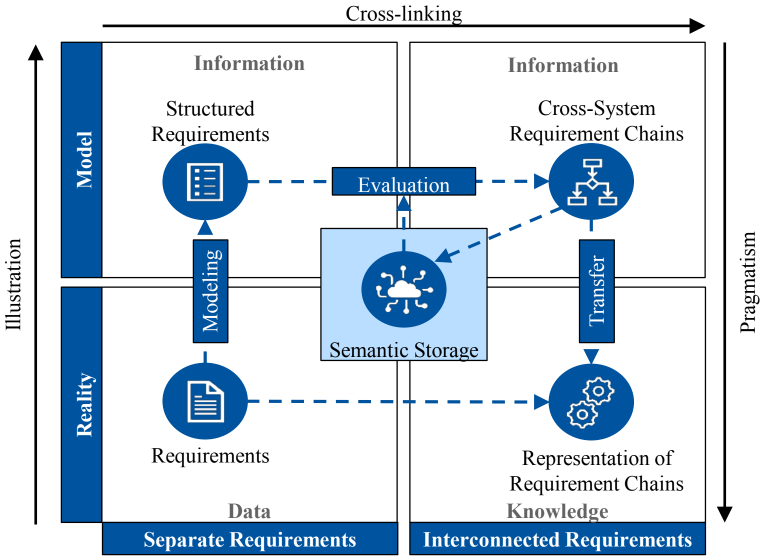


Figure 3: The process modelling framework

6. Research Gap

With help of the results of the literature research and the framework, promising approaches for solving the problem can be identified. Subsequently, a summary of the approaches and methods is given in form of a map of methods to identify gaps on the solution path and to formulate research needs based on this.

6.1 Map of Methods

The result of the literature research is an extensive landscape of methods with approaches from different areas of industry. To systematically represent the core topics, a method map is shown in Figure 4. This is intended to provide an overview of the approaches identified and to assist the reader with orientation. The systematic design of the research is reflected on the map in the form of the three snowballs NLP, AI, and modeling. The field of requirements engineering is not shown because it represents an ongoing closely linked parallel process. The Requirements field is the starting point for different paths across the map, which end in the Diagrams and Views area. To establish the reference to widespread standard solutions in the area of information technology, the complex AI solutions fastText and BERT are additionally located on the method map.

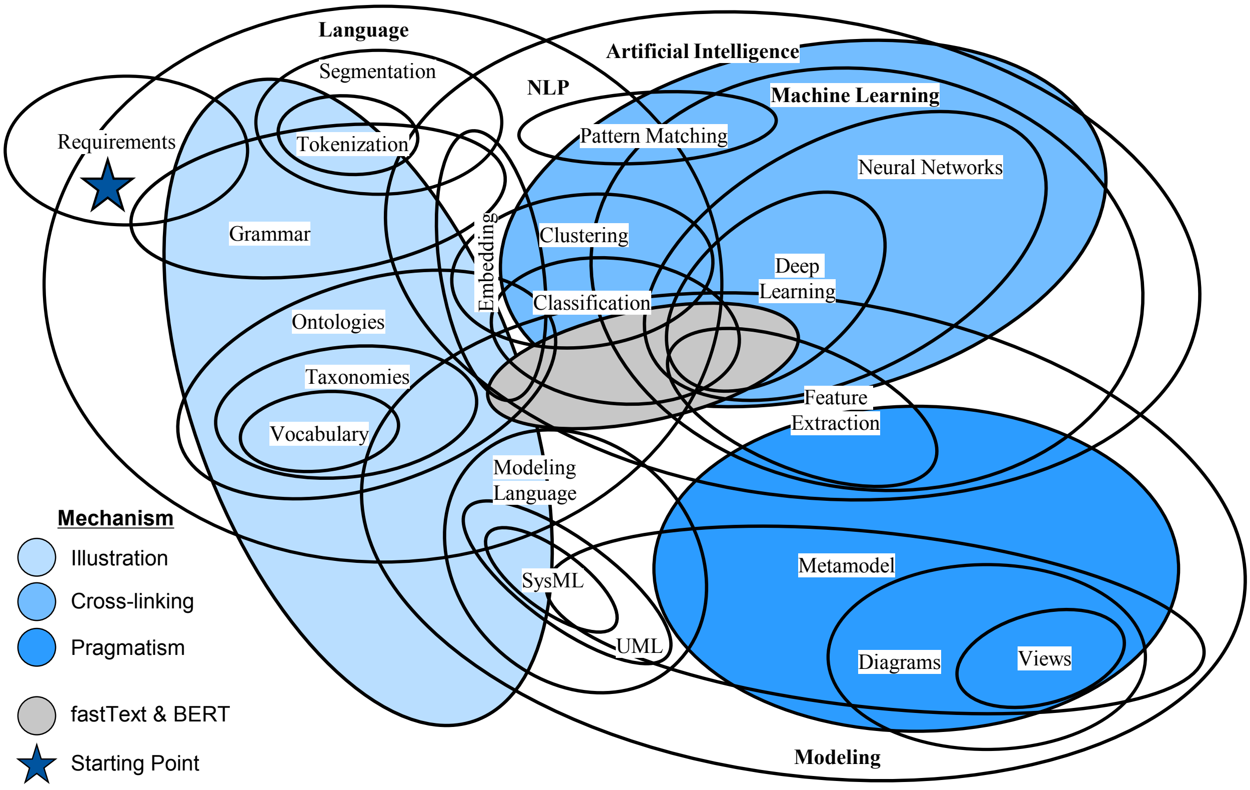


Figure 4: Map of methods

The area of illustration is particularly characterized by approaches from NLP. The systematic decomposition of requirements serves the transformation into the machine-readable domain. BASHIR ET AL., ARORA ET AL., BEN ABDESSALEM ET AL., KOCHBATI ET AL., and XU ET AL. [35][36][37][38][39] take up the methods of segmentation as well as parsing, which represent standard solutions. Based on this, ARORA ET AL. [36] describe an activity to find rule-based relationships between sentence members in a sentence. Using embeddings and similarity computation, CAMPOS ET AL. [40] and DALPIAZ ET AL. [41] filter duplicates and synonyms from the decomposed sentences. ROSADINI ET AL. [42] extends the ideas and uses parsing to predict defects in requirements. In the area of networking, activities of classification and clustering are particularly located. Classification is done by ZHOU ET AL. [43], basis for these methods provides the use of strategies from the field of neural networks and Deep Learning. ZHONG ET AL., KOWSARI ET AL., and ZHOU ET AL. [44][45][46] deal with special forms of neural networks. Different types are combined to increase efficiency and different ML methods are compared. The approaches from ZHOU ET AL. and MENG ET AL. [43][47] deal with classification based on multiple criteria. BEN ABDESSALEM ET AL. [37] and ZHONG ET AL. [44] compare the imported requirements with constructed patterns in a pattern-matching activity. According to the framework, the cross-linking of requirements takes place simultaneously with the usage of semantic information from semantic storage. While KUMAR ET AL. [48] deal with the design of ontologies from taxonomies, ZHANG ET AL. [49] use ontologies to extract semantic information. ZHOU ET AL. [46], extended ontologies by sub-ontologies. BASHIR ET AL. [35] andYANG ET AL. [50] address activities that concern the networking of requirements. To bundle the results in a final step, the design of human-readable visualizations and the application of views are mostly used in the pragmatics domain. Visualizations in the form of a Venn diagram are generated by the approach in DALPIAZ ET AL. [41]. LUCASSEN ET AL. and DALPIAZ ET AL. [51][52][41] present approaches to filter visualizations afterward and to focus or hide areas. To complement this, LUCASSEN ET AL. [51] address the highlighting of different links. Finally, YANG ET AL., ROBEER ET AL., GULIA ET AL., KOCHBATI ET AL., PEREZ ET AL. [50][53][54][37][55] present methods and activities that enable the automatic creation of diagrams.

7. Summary and Outlook

This paper presents a framework for process modeling in the AI-supported generation of chains of requirements formulated in natural language. Based on existing methods and approaches, a framework is designed that describes the intermediate states, mechanisms, and tools on the way from separate requirements to cross-system requirement chains. With the help of systematic literature research approaches from the three topic areas NLP, AI and modeling are compiled. Subsequently, representative approaches and results are presented in the form of a method map. In connection with the regulatory framework, it is thus considered which solutions can enable the selected project in a bottom-up manner. Due to the pronounced heterogeneity of the identified approaches, no explicit comparison criteria could be identified. As a result, it can be stated that a bottom-up approach based on a methodology developed in the use of existing approaches does not seem to be very effective. In future work, activities and states within the respective methods must be analyzed and, in a top-down approach, the necessary target states in the regulatory framework must first be described with sufficient precision so that the identified approaches can be compared. Especially the area of evaluation with associated networking methods from the field of neural networks is to be investigated in more detail.

References

[1] Proff, H., Fojcik, T.M., 2018. Mobilität und digitale Transformation. Springer, Wiesbaden.

[2] Schuh, G., Bergweiler, G., Fiedler, F., Slawik, V., Ahues, C., 2021. A Review Of Data-Based Methods For The Development Of An Adaptive Engineering Change System For Automotive Body Shop. Conference on Production Systems and Logistics.

[3] Eigner, M., Stelzer, R., 2013. Product Lifecycle Management: Ein Leitfaden für Product Development und Life Cycle Management. Springer, Heidelberg.

[4] Ehrlenspiel, K., 2017. Integrierte Produktentwicklung. Carl Hanser Verlag, München.

[5] Görz, G., Schmid, U., Braun, T. (Eds.), 2020. Handbuch der Künstlichen Intelligenz. De Gruyter, Berlin, Boston.

[6] Warschat, J. (Ed.), 2020. Innovation durch Natural Language Processing: Mit Künstlicher Intelligenz die Wettbewerbsfähigkeit verbessern. Carl Hanser Verlag, München.

[7] Mockenhaupt, A., 2021. Digitalisierung und Künstliche Intelligenz in der Produktion. Springer, Wiesbaden.

[8] Verein Deutscher Ingenieure. VDI 2221 - Blatt 1: Entwicklung technischer Produkte und Systeme.

[9] Rupp, C., SOPHISTen, 2014. Requirements-Engineering und -Management. Carl Hanser Verlag, München.

[10] Aunkofer, B., 2009. Anforderungsmanagement. Der Wirtschaftsingenieur. https://www.der-wirtschaftsingenieur.de/index.php/anforderungsmanagement/.

[11] Radner, W., 2000. Ein neues visuell-akustisch gestutztes Computeranalyseverfahren. Spektrum der Augenheilkunde, 239–243.

[12] Schuh, G., Stölzle, W., Straube, F., 2008. Anlaufmanagement in der Automobilindustrie erfolgreich umsetzen. Springer, Heidelberg.

[13] Hevner, A.R., 2007. A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems 19 (2), 87–92.

[14] Hevner, A.R., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS Quarterly 28 (1), 75–105.

[15] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[16] Heß, J., 2021. Regionale Erfolgsfaktoren entlang des Gründungsprozesses. Springer, Wiesbaden.

[17] Booth, A., 2006. ‘‘Brimful of STARLITE’’: toward standards for reporting literature searches. Journal of the Medical Library Association 94 (4), 421–430.

[18] Deutsches Institut für Normung e.V., 2009. DIN 69901, Projektmanagement – Projektmanagementsysteme – Teil 5: Begriffe.
[19] Ponn, J., Lindemann, U., 2011. Konzeptentwicklung und Gestaltung technischer Produkte. Springer, Heidelberg.
[20] Deutsches Institut für Normung e.V., 2015. DIN EN ISO 9001: Qualitätsmanagementsysteme.
[21] Verein Deutscher Ingenieure. VDI 2222: Methodisches Entwickeln von Lösungsprinzipien.
[22] Grande, M., 2011. 100 Minuten Für Anforderungsmanagement: Kompaktes Wissen Nicht Nur Für Projektleiter und Entwickler. Springer Vieweg, Wiesbaden.
[23] Hesse, W., Mayr, H.C., 2008. Modellierung in der Softwaretechnik: eine Bestandsaufnahme. Informatik Spektrum 31 (5), 377–393.
[24] Stachowiak, H., 1973. Allgemeine Modelltheorie. Springer, Wien.
[25] Schulze, S.-O., Schneider, A., Ackva, S. (Eds.), 2016. Erfüllung von Automotive SPICE Prozessanforderungen mittels modellbasierter Entwicklungstechniken. Carl Hanser Verlag, München.
[26] Sandmaier, H., 2019. Skalierung der physikalischen Gesetze und mathematischen Modellierung. Springer, Heidelberg.
[27] Weilkiens, T., 2006. Systems engineering mit SysML/UML. The MK/OMG Press, Heidelberg.
[28] Albers, A., Bursac, N., Scherer, H., Birk, C., Powelske, J., Muschik, S., 2019. Model-based systems engineering in modular design. Des. Sci. 5.
[29] Alt, O., 2012. Modellbasierte Systementwicklung mit SysML. Carl Hanser Verlag, München.
[30] Paaß, G., Hecker, D., 2020. Künstliche Intelligenz. Springer, Wiesbaden.
[31] Pfister, B., Kaufmann, T., 2017. Sprachverarbeitung. Springer, Heidelberg.
[32] Ertel, W., 2021. Grundkurs Künstliche Intelligenz. Springer, Wiesbaden.
[33] Heimes, H.H., 2014. Methodik zur Auswahl von Fertigungsressourcen in der Batterieproduktion. Aachen, Techn. Hochsch., Diss., 2014. Apprimus-Verl., Aachen.
[34] Maue, A., 2015. Aufwandsorientierte Gestaltung des Produktionsanlaufs am Beispiel der Automobilproduktion. Aachen, Techn. Hochsch., Diss., 2014. Apprimus-Verl., Aachen.
[35] Bashir, N., Bilal, M., Liaqat, M., Marjani, M., Malik, N., Ali, M. Modeling Class Diagram using NLP in Object-Oriented Designing. National Computing Colleges Conference, 1–6.
[36] Arora, C., Sabetzadeh, M., Briand, L., Zimmer, F., 2016. Extracting domain models from natural-language requirements. MODELS '16, 250–260.
[37] Ben Abdessalem Karaa, Wahiba, Ben Azzouz, Zeineb, Singh, A., Dey, N., S. Ashour, A., Ben Ghazala, H., 2016. Automatic builder of class diagram (ABCD): an application of UML generation from functional requirements. Softw. Pract. Exper. 46 (11), 1443–1458.
[38] Kochbati, T., Li, S., Gérard, S., Mraidha, C., 2021. From User Stories to Models: A Machine Learning Empowered Automation, 28–40.
[39] Xu, X., Cai, H., 2021. Ontology and rule-based natural language processing approach for interpreting textual regulations on underground utility infrastructure. Advanced Engineering Informatics 48.
[40] Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., Jatowt, A., 2020. YAKE! Keyword extraction from single documents using multiple local features. Information Sciences 509, 257–289.
[41] Dalpiaz, F., van der Schalk, I., Lucassen, G., 2018. Pinpointing Ambiguity and Incompleteness in Requirements Engineering via Information Visualization and NLP. Requirements Engineering: Foundation for Software Quality 10753, 119–135.
[42] Rosadini, B., Ferrari, A., Gori, G., Fantechi, A., Gnesi, S., Trotta, I., Bacherini, S., 2017. Using NLP to Detect Requirements Defects: An Industrial Experience in the Railway Domain. Requirements Engineering: Foundation for Software Quality 10153, 344–360.
[43] Zhou, P. and El-Gohary, N., 2016. Domain-Specific Hierarchical Text Classification for Supporting Automated Environmental Compliance Checking. J. Comput. Civ. Eng. (30).

[44] Zhong, B., Xing, X., Luo, H., Zhou, Q., Li, H., Rose, T., Fang, W., 2020. Deep learning-based extraction of construction procedural constraints from construction regulations. Advanced Engineering Informatics 43, 1–14.

[45] Kowsari, K., Heidarysafa, M., Brown, D.E., Meimandi, K.J., Barnes, L.E., 2018. RMDL: Random Multimodel Deep Learning for Classification. ICISDM '18, 19–28.

[46] Zhou, P., El-Gohary, N., 2016. Ontology-Based Multilabel Text Classification of Construction Regulatory Documents. J. Comput. Civ. Eng. 30 (4), 1–13.

[47] Meng, R., Zhao, S., Han, S., He, D., Brusilovsky, P., Chi, Y., 2017. Deep Keyphrase Generation. Proceedings 55th Annual Meeting of the Association for Computational Linguistics 2017, 582–592.

[48] Kumar, N., Srinathan, K., Varma, V., 2016. A graph-based unsupervised N-gram filtration technique for automatic keyphrase extraction. Int. J. Data Mining, Modelling and Management 2016 (8).

[49] Zhang, Jiansong, El-Gohary, Nora M., 2017. Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. Automation in Construction 2017 (73), 45–57.

[50] Yang, Z., Bao, Y., Yang, Y., Huang, Z., Bodeveix, J.-P., Filali, M., Gu, Z., 2021. Exploiting augmented intelligence in the modeling of safety-critical autonomous systems. Form. Asp. Comput. 33 (3), 343–384.

[51] Lucassen, G., Robeer, M., Dalpiaz, F., van der Werf, J.M.E.M., Brinkkemper, S., 2017. Extracting conceptual models from user stories with Visual Narrator. Requirements Eng 22 (3), 339–358.

[52] Lucassen, G., Dalpiaz, F., van der Werf, J.M.E.M., Brinkkemper, S., 2016. Visualizing User Story Requirements at Multiple Granularity Levels via Semantic Relatedness 9974, 463–478.

[53] Robeer, M., Lucassen, G., van der Werf, J.M.E.M., Dalpiaz, F., Brinkkemper, S., 2016. Automated Extraction of Conceptual Models from User Stories via NLP, in 2016 IEEE 24th International Requirements, pp. 196–205.

[54] Gulia, S., Choudhury, T., 2016. An efficient automated design to generate UML diagram from Natural Language Specifications. 6th International Conference - Cloud System and Big Data Engineering, 641–648.

[55] Pérez-Soler, S., Guerra, E., Lara, J. de, Jurado, F., 2017. The Rise of the (Modelling) Bots: Towards Assisted Modelling via Social Networks. IEEE/ACM International Conference on Automated Software Engineering 32, 723–728.

Acknowledgments

The authors wanted to thank the Federal Ministry for Economic Affairs and Climate Action for funding the project *KIZAM – Künstliche Intelligenz im Ziele- und Anforderungsmanagement* on basis of a decision of the German Bundestag that made the presented research possible.

Biography

Peter Burggräf (*1980) holds the Chair of International Production Engineering and Management at the University of Siegen since 2017. Together with Günther Schuh, he also heads the Factory Planning department at RWTH Aachen University and is Managing Director of StreetScooter Research GmbH.

Georg Bergweiler (*1981) is head of the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

Falko Fiedler (*1988) is group lead body shop at the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

Viktor Slawik (*1994) is a research associate at the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

Jonas Recker (*1998) is a student at the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

3rd Conference on Production Systems and Logistics

Integrating Assessment Methods In The Development Of ML-Based Business Models For Manufacturing

Felix Hoffmann[1], Enno Lang[1], Joachim Metternich[1]
[1]*Institute of Production Management, Technology and Machine Tools (PTW) / TU Darmstadt, Darmstadt, Germany*

Abstract

The use of machine learning promises great potential along the entire value chain of manufacturing companies. Many companies have already recognized the resulting opportunities for increasing enterprise value and are developing their machine learning applications for the production environment. However, despite these efforts, many of the solutions developed fail in the market. Especially small- and medium-sized enterprises have difficulties developing suitable business models for their technical applications. These difficulties arise because companies do not evaluate their business projects sufficiently during the development phases. As a result, unpromising projects are not recognized until late in the development process and thus cause high sunk costs.

This paper presents an approach for integrating assessment methods into developing machine learning-driven business models for production. Due to the diametric evolution of information availability and uncertainty during the business model development process, various methods and tools can be used for the assessment depending on the current phase. For this purpose, existing assessment methods are evaluated and contrasted regarding their suitability concerning machine learning-based business models for production. Afterwards, three approaches for the different planning phases of business model development (strategic, tactical, operational) are presented in this paper.

Keywords

Artificial Intelligence; Machine Learning; Business Models; Assessment; Manufacturing

1. Introduction

Recent studies reveal various potentials of Machine Learning (ML) for companies along the entire value chain. As a result, global GDP is expected to increase by up to 14 % or $15,7 trillion by 2030 [1]. Today, ML systems are applied across various industries. These help, for instance, to make quality management more efficient or to enable predictive maintenance of machines. Although the feasibility of these applications has already been proven many times at the research level and first marketable products are available, the actual implementation of corresponding applications and offerings lags behind the high expectations [2,3]. One of the main reasons for this discrepancy is that companies are not able to develop economically viable scenarios for their technical solutions [4]. It becomes apparent that especially those companies are successful with ML offerings that emphasize the generation of business value during development. Thus, building a business understanding and evaluating business cases as part of business model (BM) development is a key success factor for implementing competitive ML applications in manufacturing companies [5].

Research on ML-based BMs for manufacturing and the accompanying empowerment of companies is still in its infancy. The following work aims to identify appropriate evaluation methods for the different phases of maturity in the development of ML-based BMs for manufacturing. Using the tools presented, companies

DOI: https://doi.org/10.15488/12120

can evaluate the current development status of their BM and be supported in their decision-making. The underlying research hypothesis states that an appropriate methodology can be developed by analyzing existing concepts and approaches in the respective technical literature. This work considers ML as a system that "uses data, analysis and observations to perform certain tasks without needing to be programmed to do so" [6]. A BM is ML-driven if at least one relevant dimension is characterized by the use of ML methods [7].

The procedure outlined in the paper is as follows: Section 2 first provides an overview of the development process of ML-based BMs and its three-phase structure. To this end, different process models are discussed and the process is explained using a selected model. Subsequently, phase-specific requirements for the selection of evaluation methods are derived from literature. The resulting evaluation criteria are then used to evaluate existing evaluation methods which were identified in the course of a structured literature analysis. Based on this overview, section 3 proceeds with the presentation of a developed methodology, thereby choosing the most appropriate methods and tools for each planning horizon and adapting these to the area of manufacturing. For validation purposes, the introduced methodology is then applied to an actual use case from the manufacturing industry in section 4. Finally, section 5 ends with the conclusion and potential impulses for further research.

2. Research results

2.1 Development process of ML-based BMs

A successful implementation of ML-based BMs for manufacturing requires a systematic and structured process. This is of particular importance for companies that have only limited experience in the context of developing ML-based applications [3]. To address this problem, numerous process models exist in literature. Many of these approaches originate from the field of data mining, which is the extraction of structures and patterns from large amounts of data using specific analysis techniques [8]. Well-known approaches in this field include the Cross Industry Standard Process for Data Mining (CRISP-DM), the Sample, Explore, Modify, Model, Assess (SEMMA) and the Knowledge Discovery in Databases (KDD) [9]. A deeper analysis of the models in terms of their suitability for the manufacturing industry reveals numerous shortcomings that prevent their practical and holistic application in such a domain. Among the main criticisms is that existing models do not cover the process of selecting a problem and deciding whether to use ML to solve it and do not take into account the specific requirements of production environments [10].

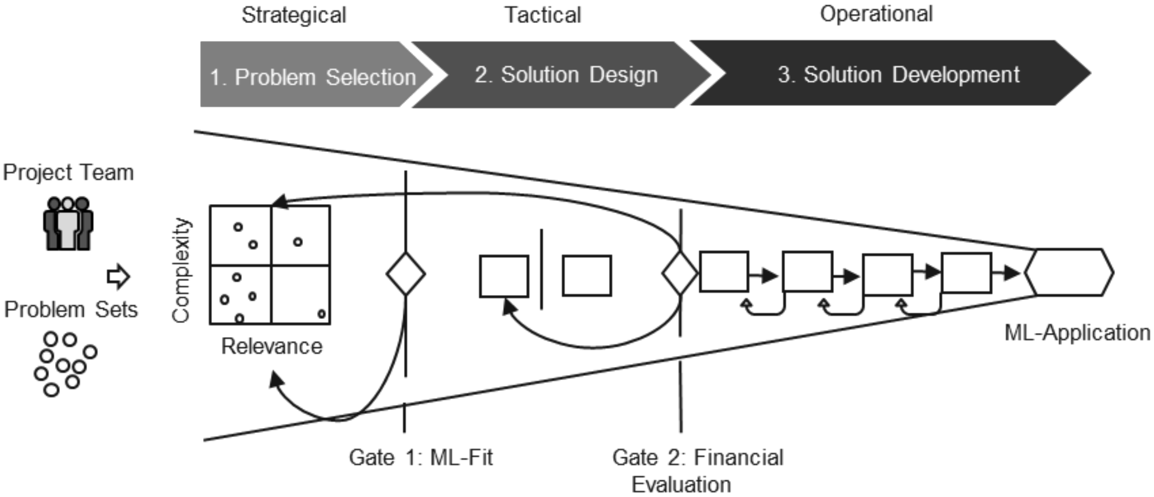


Figure 1: Development Process of ML-based BM according to Biegel et al.

In order to address these shortcomings Biegel et al. [10] introduced their AI Management Model for the Manufacturing Industry (AIMM) (see Figure 1). It is further utilized as a framework for the integration of assessment methods and to emphasize the characteristics of each phase. The process is funneled and starts with potential problems, which are subsequently transformed into an ML application using three phases: problem selection (strategic phase), solution design (tactical phase) and solution development (operational phase). A significant difference between the process phases results from the availability of relevant information and the degree of uncertainty regarding the economic viability. Relevant information includes not only forecasts on technical prospects of success, but also market scenarios and possible legal restrictions. Only a little, primarily qualitative information with a low resolution, is available at the beginning of the process, leading to a high degree of uncertainty. This uncertainty is reduced throughout the process by acquiring new and higher-quality information, gradually enabling a quantitative evaluation of projects [11].

The AIMM is designed to fail quickly in the case of an unpromising endeavor. The model takes into account that, particularly at the beginning of an application development process, the costs incurred are still low. At the same time a strong influence can be exerted on the future cost-benefit ratio in later phases of development and utilization [12]. This effect is especially relevant for the domain of production, as physical products - e.g., in the form of machine tools - are frequently linked with digital services in so-called product service systems [13]. In the case of a mere development of digital services, the share of cost emergence in early phases is proportionally higher. In contrast preventing sunk costs in later phases nevertheless has a significant impact [14,15]. Therefore, at the transitions of the phases, the AIMM process enforces to evaluate whether a problem can potentially be solved using AI technology and whether a resulting business case is financially feasible. If an idea is dropped out, the process can be revisited with a different problem, or the solution design can be adjusted accordingly. In this way, the waste of entrepreneurial resources is prevented at an early stage of the use case development [10].

A shortcoming in the AIMM is that the authors only provide a few concrete hints to the phase-specific use of assessment methods. In their approach to technology assessment from 2011, Haag et al. [16] become more specific and propose different assessment methods for the distinct process phases of technology development. However, the approach is highly technology-unspecific and does not incorporate the special requirements that arise when considering ML-based technologies, which will be highlighted throughout this paper. Due to the age of the explanations, many context-specific assessment methods (e.g., from the field of digitalization and Industry 4.0) developed in the meantime are also not included in the approach. In this regard, the method presented in this paper picks up and presents ML-suitable assessment methods in each of the three process phases.

2.2 Phase-specific requirements and evaluation of existing assessment methods

Suitable evaluation criteria must be defined for a comparable and objective evaluation of existing approaches. These result from the requirements of the different process phases and were identified as part of a structured literature review. Next is examined which activities are carried out in each phase and which input and output states are present. In addition, it is included which incoming information is available and which outgoing information must be provided for decision-making (see Figure 2).

In the **strategic phase** of problem selection, the project team first identifies and evaluates relevant problems from the production environment. As incoming information, a selection of possible problem definitions is available, whose potentials and challenges are assessed concerning a possible solution development. The underlying problem set can originate from the documentation of a continuous improvement process or from a dialogue with customers [17]. Since there can be larger problem sets, it is necessary to identify and prioritize the most promising ones. In their model, Biegel et al. propose a qualitative evaluation of problems in the two overarching dimensions of relevance and complexity using a portfolio matrix. While the number of actors and objects involved and their relationships to each other and the required interdisciplinarity is a major influencing factor for complexity, the influence on a company's key performance indicators (e.g., Overall Equipment Effectiveness) is decisive for relevance. After prioritizing the alternatives, the selected problems cross the first gate where the ML fit is examined. In this process, it is checked whether the problem under consideration meets the basic requirements for being solved with ML methods [18]. After passing through the strategic phase, quantitatively evaluated, prioritized, and ML-suitable problems remain for further pursuit in the tactical phase.

Following the results of a structured literature review, the first requirement is that suitable methods for the strategical phase are capable of enabling a comprehensible and systematic ranking of alternatives [19]. Furthermore, especially in this initial phase, it is necessary that the alternatives to be evaluated allow a *holistic assessment*, despite the low level of information [20]. In this context, it should also be possible to conduct *risk assessments* and to make *prognoses* by considering volatilities in technical, financial and organizational conditions [21]. Finally, regarding the *usability* in the strategic phase, it should be ensured that the models are generalizable to enable the evaluation of a wide range of possible problems. In addition, they should be able to determine reproducible results that are comprehensible in terms of *transparency*, even in spite of fluctuating information quality [22].

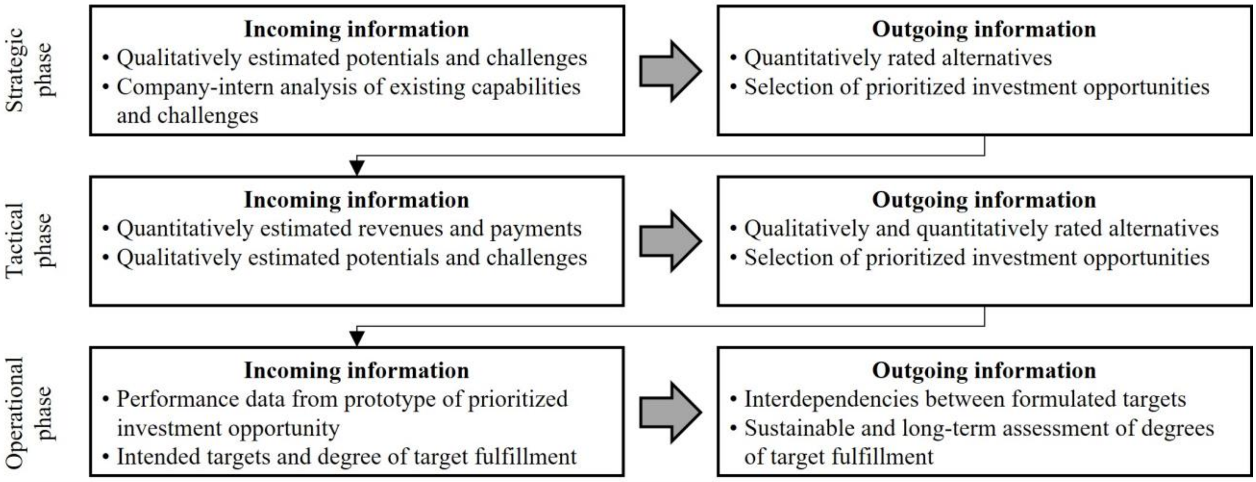


Figure 2: Overview of incoming and outgoing information in the process phases [10]

In the **tactical phase**, a selected problem is examined in more detail and developed into a draft solution in the form of a possible BM. In addition to the qualitative information already available, initial quantitative information is also available as input variables. In-depth examination of the selected problems allows estimations for possible expenses and income streams connected to implementing the respective BM. Among other things, this data can be retrieved from historical data, e.g., from service or sales or from past projects [17]. In their approach, Biegel et al. propose a financial evaluation of the project as a final gate before the solution development phase. However, in view of the extensive planning activities and the increasing availability of information, it is reasonable to further include qualitative aspects. The tactical phase thus represents a transition between the strategic phase, which is driven by qualitative information, and the numbers-driven operational phase. Therefore, suitable models for this phase must be able to merge *qualitative and quantitative aspects* and combine them in a reproducible result. However, in contrast to the previous phase, the tactical phase does not evaluate a broad problem set but various options for the BM design. This includes decisions regarding the scope of the BM (e.g., detection of errors vs. prediction of failures) or the acquisition of competencies (e.g., build up in-house vs. buy in externally) [23]. In order to weigh between these design options, the possibility of *ranking of alternatives* remains relevant in this phase.

Table 1: Evaluation of existing methods

Evaluation scale ○ not fulfilled ◐ partially fulfilled ● completely fulfilled		Data			Methodology			Practical application		
Overall systematics / focus	Evaluation model	Holistic assessment	Realistic depiction	Qualitative and quantitative aspects	Prognosis	Risk assessment	Ranking of alternatives	Usability	Transparency	Compatibility with digital BMs
		A1	A2	A3	B1	B2	B3	C1	C2	C3
Strategic	Balanced scorecard [24]	●	◐	▨	◐	◐	○	●	●	◐
	Guideline for industry 4.0 [25]	●	●	▨	◐	○	◐	●	●	●
	Canvas scoreboard [22]	◐	◐	▨	◐	◐	○	●	●	◐
	Scenario analysis [26]	●	●	▨	●	●	◐	◐	●	●
	Company culture portfolio [27]	○	●	▨	○	○	◐	●	●	◐
	Competency portfolio [28]	○	●	▨	◐	○	◐	●	●	◐
	Technology portfolio [29]	○	●	▨	○	○	◐	●	●	◐
	Growth-share matrix [30]	○	●	▨	○	○	◐	●	●	◐
	Utility analysis [31]	●	●	▨	◐	◐	●	●	◐	●
	Analytic hierarchy process (AHP) [32]	●	●	▨	◐	◐	●	◐	●	●
	Assessment model by Pokorni et al. [33]	●	●	▨	○	◐	●	◐	●	●
	Gap analysis [34]	◐	●	▨	●	○	○	◐	◐	◐
Tactical	Digitalization scorecard [35]	●	◐	●	◐	◐	○	◐	●	●
	Potential analysis [36]	◐	●	●	◐	○	◐	●	●	◐
	Assessment model by Schuh et al. [37]	●	●	●	◐	◐	●	●	●	●
Operational	Static payback method [38]	○	◐	▨	◐	○	▨	●	●	●
	Net present value method [39]	○	●	▨	◐	○	▨	●	●	●
	Dynamic payback method [38,39]	○	●	▨	◐	○	▨	◐	●	●
	Indicator system [24]	●	●	▨	○	○	▨	●	●	●
	Value driver tree [40]	◐	●	▨	●	◐	▨	◐	●	●
	Industry 4.0 maturity index [41]	●	●	▨	◐	○	▨	◐	●	●

The previously designed BM is developed and implemented as an iterative development project in the final **operational phase**. The input information in this phase consists of target values selected in the planning process, their degree of fulfilment and ML-specific performance data. The latter result from prototypical setups and testing within the development process. Thus, the use of quantitative evaluation methods is particularly indicated in the context of solution development. In their model, Biegel et al. do not specify an approach for the evaluation of the project in this phase. Nevertheless, a continuous evaluation of the project is of particular importance, especially in this resource-intensive phase [14,15]. Due to the highly dynamic nature of the iterative development process, special *usability* requirements arise in this phase. Therefore, suitable methods must enable the integration of knowledge gained from the development process in the short term. Furthermore, the results of the applied evaluation methods must be comprehensible for all stakeholders involved in the project in terms of *transparency*. Finally, especially in the solution development phase, there is a high demand for the evaluation methods regarding their *compatibility with digital business models*. ML-specific figures must be considered to a greater extent, especially regarding possible optimizations by the underlying application. For example, the expected prediction accuracy of the model is decisive for the profitability of an ML use case.

Table 1 shows the results of an assessment of existing evaluation methods using the derived evaluation criteria. The evaluation methods were identified through systematic literature analysis. The methods are divided according to the three process phases: strategic, tactical, and operational. The evaluation criteria are divided into three areas: data, methodology and practical application. The data area includes the criteria *holistic assessment*, inclusion of *qualitative and quantitative aspects* as well as *realistic depiction*. The criterion "realistic depiction", which has not been mentioned so far, refers to the fact that the recording and preparation of all necessary technical and business contexts is necessary for a well-founded evaluation. It is thus relevant for all process phases [42].

3. Description of the assessment methodology

In the following, the developed methodology is introduced. The selection and combination of methods is based on the evaluation of existing approaches which was presented in chapter 2. An overview of the methodology is shown in Figure 3.

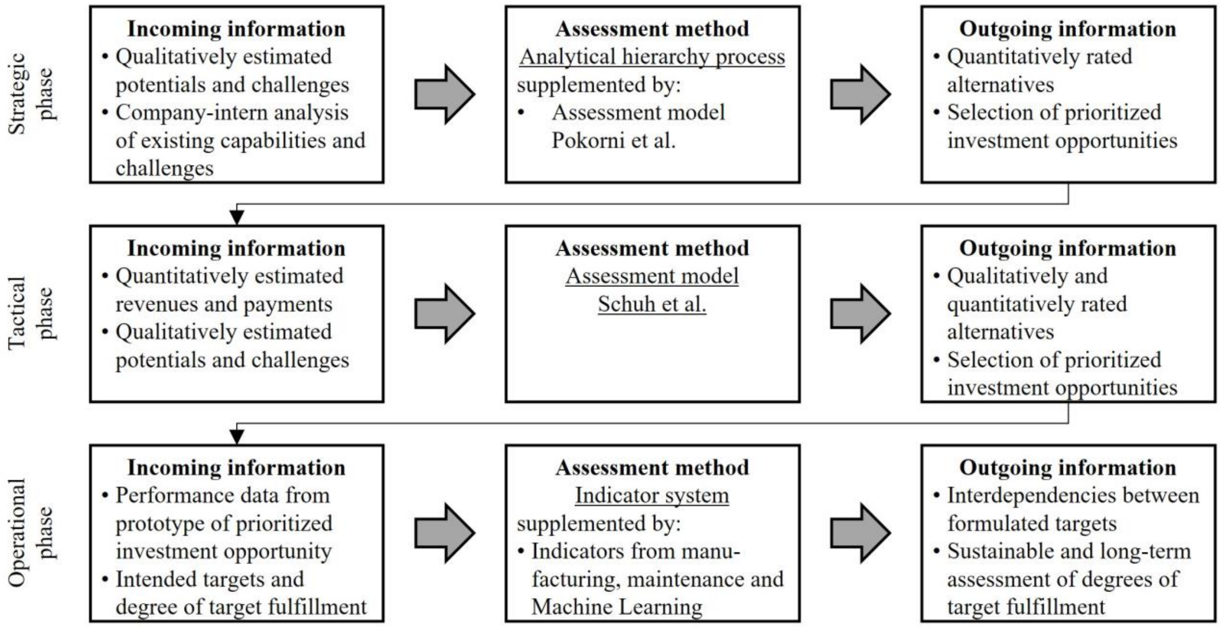


Figure 3: Three-levelled assessment methodology

For the **strategic phase**, the analytic hierarchy process provides the main structure and is supplemented by aspects from the assessment model according to Pokorni et al. In this phase, it is important that a holistic evaluation of the problem set can take place and subsequently a ranking of alternatives is made possible. Among the approaches examined, the utility analysis, the analytic hierarchy process (AHP) and the assessment model by Pokorni et al. fulfil these requirements at the best. Comparing the utility analysis and the AHP, there is a decisive difference: the AHP has an iterative structure and provides a consistency check to avoid logic errors. On the one hand, this makes the AHP more transparent than the utility analysis due to the systematic assurance of consistency. On the other hand, the usability deteriorates due to the increased effort required to perform the analysis. However, this additional effort in the AHP is reduced by a computer-aided execution of the procedure [38]. Nevertheless, due to the high degree of generalizability of the AHP, it makes sense to enrich the procedure with elements from the industry- and context-specific approach of Pokorni et al. In this model, positive effects of the use of ML are interpreted as potentials (e.g., increase in efficiency, increase in productivity) and negative effects as challenges (e.g., implementation costs, compliance challenges).

Due to its holistic approach and good performance overall, the assessment model according to Schuh et al. [37] is selected and adapted for the **tactical phase**. The hybrid model focuses on an assessment of effort and benefit, considering challenges and fields of action of ML-driven BMs. As the digitization scorecard and the potential analysis, this method is suitable for the required use of qualitative and quantitative data. Compared to other approaches, the method of Schuh et al. stands out due to its possibility of ranking alternatives. Within the model, the evaluation of qualitative input information merges with the results of quantitative evaluations in a portfolio matrix. By using and adapting an indifference curve within the model, it is also possible to include user-specific preferences. The position to the indifference curve is used to decide whether a solution design is perceived as an investment decision and transferred to the effort-intensive operational phase. Accordingly, adjustments can be made or an use case is discarded completely [37].

Finally, for the **operational phase**, an indicator system is introduced. It combines domain-specific key figures from the field of manufacturing with ML-based key figures. Thereby, it aims at enabling a sustainable assessment and control of the BM during operation. Compared to methods of investment calculation (static/dynamic payback method; net present value method), an indicator system is more suitable for a holistic assessment. Thus, in addition to purely financial key figures, ML-specific (e.g., precision, accuracy) and production-specific key figures (e.g., utilization rate, productivity) can be included. Compared to the value driver tree, the indicator system distinguishes itself by better usability. Accordingly, it is possible to draw on metrics already known and used by the various stakeholders in the interdisciplinary development project. Finally, the effort of the project team to create the system consists of identifying dependencies between individual key figures and linking them. The remaining Industry 4.0 maturity index is also partially based on a system of key performance indicators. Therefore, the evaluation method of the operational phase also takes up aspects from the model's key figure system.

4. Application and validation

To validate the presented approach, the procedure introduced in this paper was applied to an exemplary use case. It originates from the industry-centric research project "Sensorische Schutzabdeckung" which was funded by the LOEWE – State offensive in Hessen, Germany. The basic idea of this project was to develop a predictive maintenance application for protective covers in machine tools which was realized by applying ML to gathered retro-fit sensor data [43]. In the following, the application of the approach presented in chapter 4 is applied to the use case.

Compared to an application in a real industrial environment, there is a significant difference when applied to a research project: Whereas in industry one often must choose between working on different problems arising from one's own company or from customer requirements, the problem in a research project is usually already defined in advance. For this reason, it was decided not to apply the assessment method from the strategic phase and to start with the application of the methodology from the tactical phase. It was possible to consider both the potentials of an ML-based use case compared to the previous status quo in the maintenance of protective covers and the challenges that exist along the way. The challenges were incorporated into the evaluation process as qualitative aspects. The biggest challenge identified was that the company itself had little experience in the field of data analysis and that the sensor technology required for data acquisition in the ML application had not existed as a ready-made solution. In this way, important key partners for future development activities could be identified and acquired. At the same time, it was possible to use extensive quantitative input. Thus, a potential cost saving for the avoidance of too early or too late repair measures as well as a customer's willingness to pay and possible unit number ranges for a marketable solution could be determined and included in the evaluation. In doing so, the application of the evaluation method resulted in a positive prognosis for a decision to invest in the development of the ML-based product-service system.

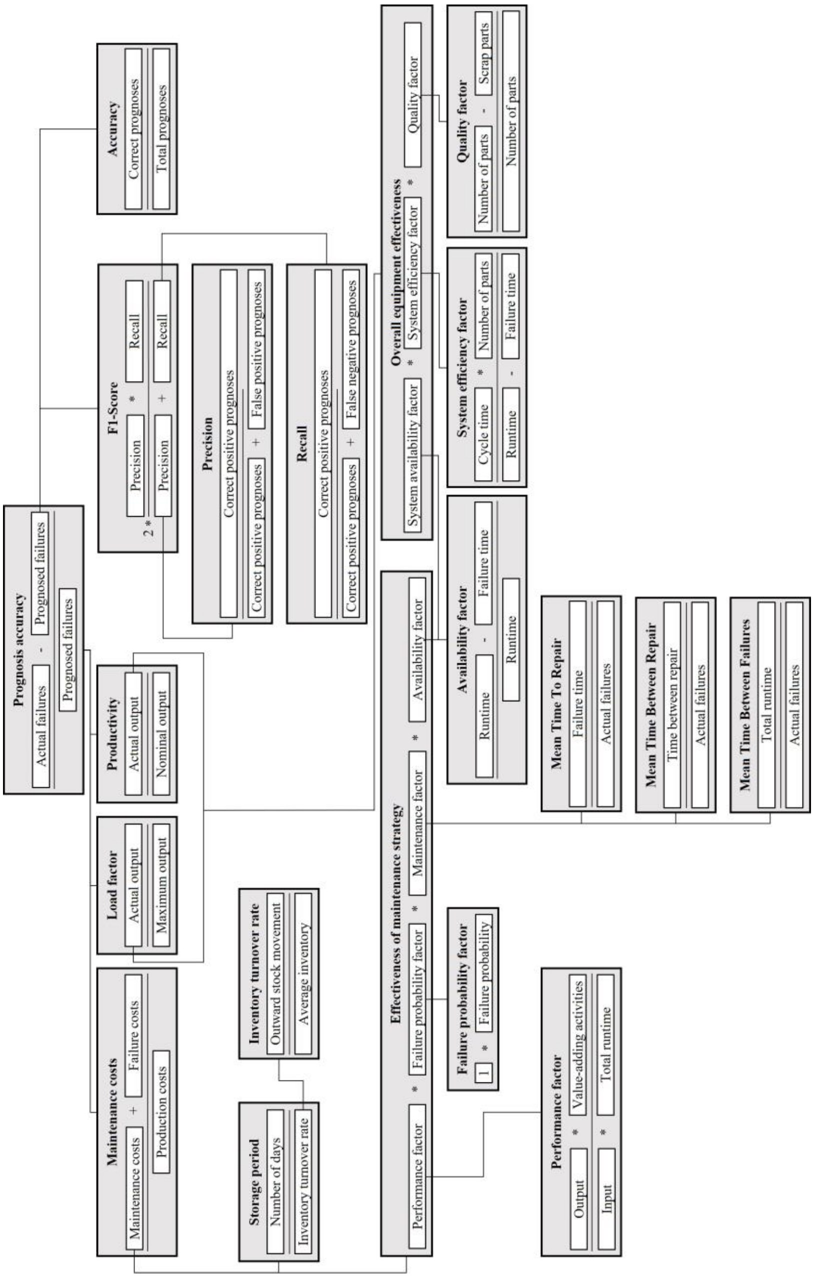


Figure 4: Indicator system for operational phase in the project “Sensorische Schutzabdeckung”

In the subsequent phases, the evaluation method for the operational phase was applied (see Figure 4). Within the resulting indicator system, various context-related key figures are related to each other via mathematical and logical connections. The key figures are taken from both the business management and ML area. In addition, specific key figures from the field of maintenance are included, such as the mean time to repair or the failure probability factor. During the development of the indicator system, the stakeholders involved - namely the management, the company's domain experts, sensor and electronics developers and data scientists - were thus able to incorporate their relevant indicators. This confirmed a good usability of the approach within the framework of an interdisciplinary project team. By establishing and tracking the indicators, important levers for achieving an economically viable scenario were identified during the course of the project. As a result, it was possible, among other things, to substitute electronic components with less expensive variants and to reduce the amount of data processed.

Within the model shown in Figure 4, mathematical relationships are highlighted by operators, whereas logical relationships are indicated by connecting lines. The data used in the project originated from different sources. Business- and maintenance-specific data was already available through previous research from the company's sales and service departments. ML-specific data was collected as part of the development project and the tracked metrics in the model were subject to significant changes. During development, various sensor and ML concepts were designed and investigated, prototypes were built and experiments were conducted. This enabled more precise figures to be derived for possible model qualities and estimates to be made of the hardware required for implementation. Here, the indicator system confirmed its advantage of incorporating new findings within a short-cycle development process. Since the research project did not comprise a complete product development but ended as an extended feasibility study, reliable figures were not yet available for all aspects at this time. Nevertheless, offers from external contractors and a more in-market analysis were acquired at the end of the project. Thus, it was possible to use the assessment methodology to draw up possible scenarios a continuation of development efforts within the company.

5. Conclusion and future research

Given the increasing availability of data, the importance of ML and its integration into companies' use cases and BMs is higher than ever before. Assessment methods are meant to evaluate the profitability and viability of BMs during their development, therefore intending to reduce sunk costs by prioritizing promising alternatives in the early stages of development. However, due to the variety of methods and novel potentials as well as challenges coming with ML-based BMs, companies still struggle to find appropriate ways of assessing their BMs. In this paper, a three-levelled methodology for assessing ML-based BMs has been introduced. Considering the strategic, tactical, and operational planning horizons, various assessment methods have been assigned, rated, combined and adapted into a holistic assessment methodology for ML-based BMs. Following the depicted gates within the AIMM from Biegel et al., the developed methodology allows less promising alternatives to drop out and thus to reduce sunk costs. Furthermore, the indicator system, containing of business-, application- and ML-related indicators, enables continuous tracking of BMs after being implemented in practice.

The introduced methodology has been validated using an actual use case from the manufacturing industry. However, given the novelty of the approach, further validations, especially within the strategical phase of the methodology, are necessary – some have already been initiated. Furthermore, in the context of this paper, a broad overview of the selected methods was provided based on their theoretical evaluation. In the future, it is necessary to describe the developed methodology and especially its associated methods in more detail and give further instructions for the practical implementation and application.

6. Acknowledgements

The project KompAKI is financed with funding provided by the German Federal Ministry of Education and Research within the “Zukunft der Wertschöpfung – Forschung für Produktion, Dienstleistung und Arbeit” program under project funding reference number 02L19C150. It is implemented by the PTKA Project Management Agency. The author is responsible for the content of this publication.

7. Literature

[1] Rao, A., Verweij, G., 2017. Sizing the prize: What’s the real value of AI for your business and how can you capitalise? https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf. Accessed 7 December 2021.

[2] Berg, A., 2019. Industrie 4.0: Jetzt mit KI. bitkom, Hannover. https://www.bitkom.org/sites/default/files/2019-04/bitkom-pressekonferenz_industrie_4.0_01_04_2019_prasentation_0.pdf. Accessed 6 December 2021.

[3] Metternich, J., Biegel, T., Bretones Cassoli, B., Hoffmann, F., Jourdan, N., Rosemeyer, J., Stanula, P., Ziegenbein, A., 2021. Künstliche Intelligenz zur Umsetzung von Industrie 4.0 im Mittelstand: Expertise des Forschungsbeirats der Plattform Industrie 4.0, München. https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/Expertise-Forschungsbeirat_KI-fuer-Industrie40.pdf?__blob=publicationFile&v=3. Accessed 6 December 2021.

[4] Lundborg, M., Märkel, C., 2019. Künstliche Intelligenz im Mittelstand: Relevanz, Anwendungen, Transfer. WIK GmbH, Bad Honnef.

[5] Ermakova, T., Blume, J., Fabian, B., Fomenko, E., Berlin, M., Hauswirth, M., 2021. Beyond the Hype: Why Do Data-Driven Projects Fail?, in: Proceedings of the 54th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences.

[6] Antonescu, M., Calitatea, B., 2018. Are business leaders prepared to handle the upcoming revolution in business artificial intelligence?, Bucharest.

[7] Bretones Cassoli, B., Hoffmann, F., Metternich, J., 2021. Comparison of AI-Based Business Models in Manufacturing: Case Studies on Predictive Maintenance.

[8] Hand, D.J., Adams, N.M., 2014. Data Mining, in: Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L. (Eds.), Wiley StatsRef: Statistics Reference Online, vol. 5. Wiley, pp. 1–7.

[9] Azeved, A., Rojao, I., Santos, M. (Eds.), 2008. KDD, SEMMA and CRISP-DM: A parallel overview.

[10] Biegel, T., Bretones Cassoli, B., Hoffmann, F., Jourdan, N., Metternich, J., 2021. An AI Management Model for the Manufacturing Industry - AIMM.

[11] Samset, K., Volden, G.H., 2016. Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. International Journal of Project Management 34 (2), 297–313.

[12] Ehrlenspiel, K., Kiewert, A., Lindemann, U., 1998. Kostenverantwortung der Produktentwickler, in: Ehrlenspiel, K., Kiewert, A., Lindemann, U. (Eds.), Kostengünstig Entwickeln und Konstruieren. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 5–16.

[13] Baines, T.S., Lightfoot, H.W., Evans, S., Neely, A., Greenough, R., Peppard, J., Roy, R., Shehab, E., Braganza, A., Tiwari, A., Alcock, J.R., Angus, J.P., Bastl, M., Cousens, A., Irving, P., Johnson, M., Kingston, J., Lockett, H., Martinez, V., Michele, P., Tranfield, D., Walton, I.M., Wilson, H., 2007. State-of-the-art in product-service systems. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 221 (10), 1543–1552.

[14] Zelkowitz, M., Shaw, A., Gannon, J., 1979. Principles of Software Engineering and Design. Prentice-Hall, New Jersey.

[15] Schach, S.R., 1999. Software Engineering. McGraw-Hill, Boston.

[16] Haag, C., Schuh, G., Kreysa, J., Schmelter, K., 2011. Technologiebewertung, in: Schuh, G., Klappert, S. (Eds.), Technologiemanagement, vol. 10. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 309–366.

[17] Möhrle, M.G., 1999. Der richtige Projekt-Mix. Springer Berlin Heidelberg, Berlin, Heidelberg.

[18] Geschka, H., Geschka, M., 2015. Innovationsmanagement: Ideenfindung und Konzeptentwicklung. AKAD, Stuttgart.

[19] Ahsen, A. von, Heesen, M., Kuchenbuch, A., 2010. Grundlagen der Bewertung von Innovationen im Mittelstand, in: Ahsen, A. (Ed.), Bewertung von Innovationen im Mittelstand, vol. 5. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–38.

[20] Rehme, M., Lindner, R., Götze, U., 2015. Perspektiven für Geschäftsmodelle der Fahrstrombereitstellung, in: Proff, H. (Ed.), Entscheidungen beim Übergang in die Elektromobilität, vol. 4. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 409–428.

[21] Taeuscher, K., Abdelkafi, N. (Eds.), 2015. Business Model Robustness: A System Dynamics Approach.

[22] Batocchio, A., Ferraz Minatogawa, V.L., Anholon, R., 2017. Proposal for a Method for Business Model Performance Assessment: Toward an Experimentation Tool for Business Model Innovation. Journal of Technology Management & Innovation 12 (1), 61–70.

[23] Vahs, D., Brem, A., 2015. Innovationsmanagement: Von der Idee zur erfolgreichen Vermarktung. Schäffer Pöschel, Stuttgart.

[24] Gottmann, J., 2019. Produktionscontrolling: Wertströme und Kosten Optimieren, 2nd ed. ed. Gabler, Wiesbaden, 233 pp.

[25] Anderl, R., Picard, A., Wang, Y., Fleischer, J., Dosch, S., Klee, B., Bauer, J., 2015. Leitfaden Industrie 4.0: Orientierungshilfe zur Einführung in den Mittelstand. VDMA-Verl., Frankfurt am Main, CI.

[26] Tesch, J. (Ed.), 2016. Discovering the Role of Scenario Planning as an Evaluation Methodology for Business Models in the Era of the Internet of Things (IoT).

[27] Denison, D.R., Mishra, A.K., 1995. Toward a Theory of Organizational Culture and Effectiveness. Organization Science 6 (2), 204–223.

[28] Steinle, C., 2005. Ganzheitliches Management: Eine mehrdimensionale Sichtweise integrierter Unternehmungsführung. Gabler Verlag, Wiesbaden, 912 pp.

[29] Höft, U., 1992. Lebenszykluskonzepte: Grundlage für das strategische Marketing- und Technologiemanagement. Erich Schmidt Verlag GmbH & Co KG, Berlin, Germany.

[30] Øivind Madsen, D., 2017. Not dead yet: the rise, fall and persistence of the BCG Matrix. Problems and Perspectives in Management 15 (1), 19–34.

[31] Schulte, G., 2007. Investition: Investitionscontrolling und Investitionsrechnung, 2., überarb. Aufl. ed. Oldenbourg Verlag, München, 235 pp.

[32] Bhushan, N., Rai, K., 2004. Strategic Decision Making: Applying the Analytic Hierarchy Process. Springer, London.

[33] Pokorni, B., Braun, M., Knecht, C., 2021. Menschzentrierte KI-Anwendungen in der Produktion - Praxiserfahrungen und Leitfaden zu betrieblichen Einführungsstrategien, Stutgart.

[34] Kim, S., Ji, Y., 2018. Gap Analysis, in: Heath, R.L., Johansen, W. (Eds.), The International Encyclopedia of Strategic Communication. Wiley, pp. 1–6.

[35] Schuhknecht, F. Performance Management in der digitalen Welt. Dissertation. Springer Fachmedien Wiesbaden, Wiesbaden.

[36] Obermaier, R., Hofmann, J., Wagenseil, V., 2019. Systematische Abschätzung von Wirtschaftlichkeitseffekten von Industrie-4.0-Investitionen mithilfe von Prozess- und Potenzialanalysen, in: Obermaier, R. (Ed.), Handbuch Industrie 4.0 und Digitale Transformation, M64. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 189–203.

[37] Schuh, G., Boos, W., Kelzenberg, C.M.B., Lange, J. de, Stracke, F.K., Helbig, J., Boshof, J., Ebbecke, C.W., 2018. Industrie 4.0: Implement it!: Ein Leitfaden zur erfolgreichen Implementierung von Industrie 4.0-Lösungen, 1st ed. RWTH Aachen Werkzeugmaschinenlabor, Aachen.

[38] Götze, U., 2014. Investitionsrechnung. Springer Berlin Heidelberg, Berlin, Heidelberg.

[39] Schuster, T., Rüdt von Collenberg, L., 2017. Investitionsrechnung: Kapitalwert, Zinsfuß, Annuität, Amortisation. Springer Berlin Heidelberg, Berlin, Heidelberg.

[40] Kaufmann, T., Servatius, H.-G., 2020. Das Internet der Dinge und Künstliche Intelligenz als Game Changer. Springer Fachmedien Wiesbaden, Wiesbaden.

[41] Zeller, V., Hocken, C., Stich, V., 2018. acatech Industrie 4.0 Maturity Index – A Multidimensional Maturity Model, in: Moon, I., Lee, G.M., Park, J., Kiritsis, D., Cieminski, G. von (Eds.), Advances in Production Management Systems. Smart Manufacturing for Industry 4.0, vol. 536. Springer International Publishing, Cham, pp. 105–113.

[42] Eiselmayer, K., Kottbauer, M., 2015. Trends im Controlling. Controller Magazin 40, 24–25.

[43] Hoffmann, F., Brockhaus, B., Metternich, J., 2020. Predictive Maintenance für Schutzabdeckungen: Vom Geschäftsmodell zur Anwendung. WT Werkstattstechnik 110 (07-08), 496–500.References

8. Biography

Felix Hoffmann, M. Sc. is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2019. His research interests are ML-driven Business Models in manufacturing and Predictive Maintenance.

Enno Lang, M. Sc. is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2021. His research interests are ML-driven Business Models in manufacturing and data-based process optimization.

Prof. Dr.-Ing. Joachim Metternich is the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2012. In addition, Prof. Metternich is the spokesman for the Center of Excellence for Medium-Sized Businesses in Darmstadt and president of the International Association of Learning Factories.

3rd Conference on Production Systems and Logistics

Detecting Deterministic Chaotic Inter-arrival Times in Material Flow Systems

Martin Manns[1], Denny Höhnen[1]
[1]*PROTECH - Institute for Production Technology / Universität Siegen, Siegen, Germany*

Abstract

Automated, modular, asynchronous and locally controlled material flow systems promise high routing flexibility in production lines because their conveying modules can be reconfigured without reprogramming PLCs. However, if such material flow systems comprise cycles and different routes, they may exhibit undesirable deterministic chaotic inter-arrival times, which can lead to conveying bottlenecks when approaching maximum capacity. Since existing analytical models have not been practically adopted for planning material flow systems, an approach for detecting deterministic chaotic inter-arrival times during production is proposed. It employs the Hough transform to identify trajectories in inter-arrival time phase space. The approach is tested with a laboratory double belt conveyor system, in which non-deterministic behavior is minimized. Results are compared with a previously published analytical model. It is shown that the proposed approach is able to detect deterministic chaotic inter-arrival times for the test cases. Phase trajectories are only partly identified. Future research should test and compare different line detection algorithms for their influence on the approach's robustness in practical production environments.

Keywords

Assembly control; Material-flow; Discrete event systems; Time complexity; Automation

1. Introduction

In series production of piece goods, material-flow automation is a common approach to reduce non-value adding labor cost. This approach considerably affects production efficiency because the time of material handling is a considerable part of total manufacturing time. Following [1], its ratio is about 85%.

A popular class of automated material transport systems employs double conveyor belts to transport work pieces that are located on work piece holders. Such systems support routing flexibility. Two conveyor belts are normally moving at constant speed and take work piece holders with them using friction. When a work piece holder encounters an obstacle, e. g. a stopper or another work piece holder that is already blocked, it stops while the conveyor belts continue to move underneath it. When the block finishes, the obstacle is removed and the work piece holder is again moved by the belts. For such conveyor systems, distributed control can be used to make them reconfigurable [2].

A typical layout of double conveyor belts comprises a main loop and several side loops, in which stations are situated. Each work piece holder contains a memory with its production plan and a pointer to its next production step. At each junction, the memory of an arriving work piece holder is read out e. g. via RFID. If a station that is viable for conducting the next production step is situated inside the side loop then - if possible - the work piece holder is routed to this side loop. Otherwise, it continues its journey in the main loop. This design allows decentralized control of the material flow system because a Programmable Logic

DOI: https://doi.org/10.15488/12124

Controller (PLC) that is responsible for controlling one side loop does not require knowledge about the other side loops but may only communicate via the work piece holders' memories.

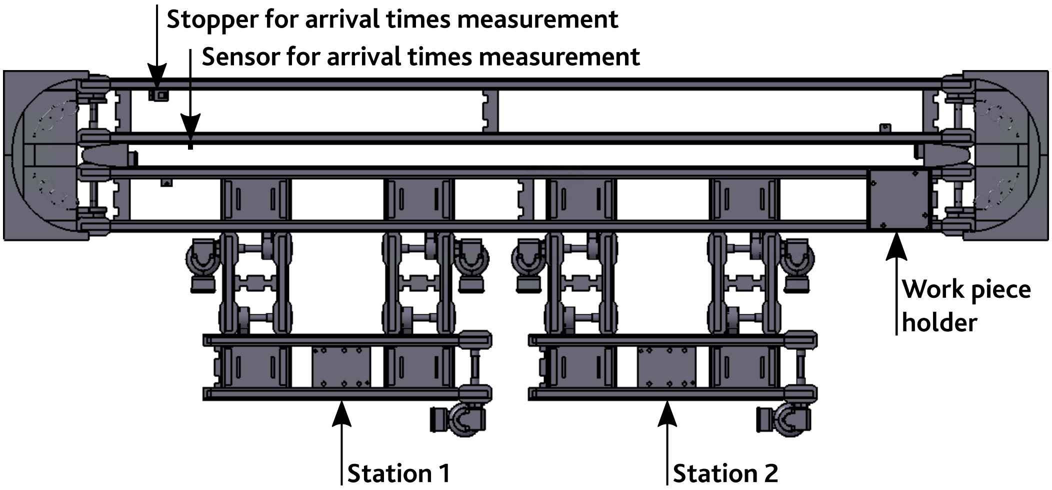


Figure 1: Double conveyor belt system with one work piece holder.

This control design allows reconfiguration of the material flow system layout without reprogramming PLCs, which can be considered an advantage for agile production (c. [2]). However, it also implies that the material flows through the different routes are normally not synchronized. It has been shown in [3] that such non-synchronized routes in material flows may lead to deterministic chaotic behavior, which may lead to reduced conveying capacity that can reduce overall production efficiency of assembly systems.

In practice, such state explosions are either ignored or they are counteracted with system designs that limit the number of system states but that are harder to reconfigure at the control level. Considering non-synchronized production systems or production system chains this approach becomes infeasible.

Relevance of chaotic behavior in production systems with loops has been reported by several researchers (c. [4] and [5]). While effects on efficiency may be dominated in systems with high down-time ratio, it can lead to conveying bottlenecks when approaching maximum capacity. Furthermore, deterministic chaotic behavior should be generally avoided because it introduces unlimited numbers of system states, i. e. dynamic complexity. In [6], it has been shown that stochastic effects such as processing time variability overlay but do not cancel this effect. Therefore, it would be desirable to detect such deterministic chaotic behavior when it happens during production in order to be able to counteract it e. g. by re-initialization. In [4], a mathematical model of autonomously controlled production networks considering time delay systems is described, which allows stability analysis using Lyapunov functions. However, practical management or control of the effect are considered hard because published models of the effect in [3] or [4] are difficult to understand and handle in engineering practice.

After a short overview of present approaches to material flow analysis, this work presents a novel, practically applicable method for automatic detection of deterministic-chaotic material flow behavior. The method provides a measure that describes the complexity of the trajectory that the inter-arrival times converge to in phase space. In this context, the term phase space is considered in the context of discrete event systems and shall be defined as a space of two dimensional vectors, where the first component is the n-th and the second component is the (n+1)-th inter-arrival time. In a proof of principle, the method is tested using a double belt conveyor system with two work stations that are situated in separate side flows. Stochastic behavior is minimized in the tests. Should the method prove applicable to real world scenarios, it could be used for on-line detection of chaotic material flow behavior in production systems.

2. Models of deterministic chaotic material flow behavior

2.1 Material flow simulation

Today, material flow simulation is an established technology that is widely adopted throughout industry (c. [7]). The dominant paradigm is to employ discrete event simulation models that are set up and parameterized in graphical user interfaces. Graphical definition of material flows and control flows are often accompanied by code that describes behavioral logic. An approach for reducing manual modeling effort has been proposed in [8].

Material flow simulation has been employed to investigate deterministic chaotic effects with limitless growing state space in [6]. It was found that deterministic chaotic effects overlay stochastic effect so that either may be dominant (s. Figure 2).

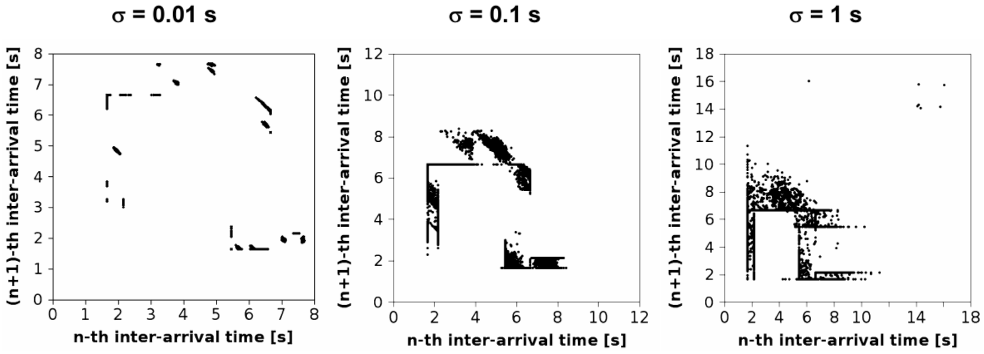


Figure 2: Simulated inter-arrival time patterns for normal distributed bottleneck station processing times with different standard deviations for a double belt conveyor belt system using 57 work piece holders [6].

A different approach is to employ physics simulation for material flow simulations to reduce modeling effort for the non-controlled environment behavior (s. [9]). Resulting motions are less abstracted than those of discrete event simulations. Therefore, accelerations can be investigated, which enables optimizing conveying velocities (c. [9]). [10] provides a survey of 3D game engines that comprise physics considering their application for production system simulation.

2.2 Analytical material flow models

A well researched domain of analytical approaches to model and analyze material flow behavior is queueing theory (s. [11]). There, arrival times are assumed to follow stochastic distributions. If complex layouts are considered, these distributions normally follow specific types such as exponential distributions (c. [12]).

Furthermore, colored Petri-networks (c. [13]) and max-plus dioids (c. [14]) have been employed for material-flow analysis of flexible manufacturing systems. Typically, in these analytical approaches deterministic chaotic behavior is ruled out by a-priori assumptions or by modeling rules that are required for applying model analyses.

In [3], series based analytical models for deterministic chaotic inter-arrival time behavior has been presented. However, practical applications of the approach are limited because an analytical measure for comparing or assessing the time series is missing.

3. Method for detecting deterministic chaotic inter arrival times

The main idea of the method is to treat scatter plots of the phase space of inter-arrival times, i. e. the plot of the n-th and (n+1)-th inter-arrival time as images, on which the probabilistic Hough line transform [15] is employed for line detection. If lines are detected then there are visible trajectories in phase space that indicate deterministic chaotic behavior. The number of lines is employed as an indicator of dynamic complexity.

As an input, the proposed method employs arrival times of work piece holders at a specific location in the main loop. This location is situated in the region in front of a fork joint that leads to a bottleneck station. In front of the location, there must be a queueable conveyor section (s. Figure 1). Since modern PLCs have access to an internal clock, detection of arrival times is easily implementable. However, for the method time accuracy is crucial so that the logging of arrival times should run in a fast loop that is separated from the main control program. The method's steps are conducted as follows:

- Arrival times are stored with a sampling frequency of 1000 Hz in a ring buffer of 100 values inside the PLC.
- Each 20 ms, the ring buffer is read out by a computer that is connected via field bus.
- After each read out, the new arrival times are converted into inter-arrival times by subtracting the previous arrival time.
- Each inter-arrival time is mapped to a 2D histogram with 100 bins for each axis (optimum bin size may vary for different systems), in which the x-coordinate represents the (n)-th inter-arrival time and the y-coordinate represents the (n+1)-th inter-arrival time.
- A binary matrix is set up, in which each element corresponds to a bin. Each matrix element is set to 255 for empty bins and to 0 for non-empty bins.
- The matrix is extended by 1/5 its size at each border, i. e. 20 elements are padded at the top, bottom, left and right border so that the matrix size becomes 140x140.
- The Canny edge detector [16] is applied to the matrix as it were an image.
- The probabilistic Hough line transform (s. [15]) is applied on the result.
- Resulting lines are counted.

Table 1 provides an overview of the parameters for the Canny edge detector and the probabilistic Hough line transform, which have been manually derived.

Table 1: Parameters for the algorithms

Algorithm	Parameter	Value
Canny	Gaussian filter kernel size (x and y)	10
	Hysteresis procedure threshold 1	100
	Hysteresis procedure threshold 2	200
	Sobel operator aperture size	7
Probabilistic Hough Line Transform	Distance resolution	20
	Angle resolution [rad]	0.17
	Voting threshold	50
	Maximum gap between points	30

4. Tests setup

The tests have been conducted with a Bosch TS/2+ based conveyor system (s. Figure 3). The double belt conveyor system comprises a main loop and two side loops. Processing times of both stations are 10 s. Conveying speed is 0.22 m/s. The production plan for each work piece holder is Station 1 → Station 2 → Station 1 → etc.

Figure 3: Test setup with double belt conveyor

The overall system is controlled by three independently operating PLCs (one for the main loop and one for each side loop). The main control program cycles of the three PLCs are not synchronized with each other. Arrival times measurements are performed independently on one PLC so that the PLC runs both the main control program and a measurement program with 1 ms cycle time.

Behavior at the joints follows one pattern that is commonly found in industry:

When a work piece holder arrives at a fork joint of a side loop with its next station then it is immediately routed into the side loop if possible. If the side loop is blocked because of a queue in front of the bottleneck station, the work piece holder waits for 4 s, during which it is routed to the side loop as soon as no block is present any more. If the block lasts longer then the work piece holder continues traveling along the main loop.

When a work piece holder arrives at a merge joint, the side loop always gets precedence so that blocking of the station inside the side loop is avoided.

Work piece holders are initially queued at the stopper behind the long outer section of the main loop (front conveyor right in front of the turning unit in Figure 3)

Tests are conducted with each number of work piece holder numbers starting at 1 and ending at 32. At 32 work piece holders, a deadlock immediately occurs. Note that deadlock situations start occurring at 26 work piece holders and above after less than one minute.

5. Results

For 1 to 7 work piece holders, all detected lines are situated close to the x and y axes, and no chaotic behavior can be observed. Note that for tests with 3 work piece holders or less, no lines are detected because of the maximum gap parameter employed.

Figure 4 visualizes the test results for 7 to 14 work piece holders. Doing a visual analysis, an increasingly relevant pattern that covers areas far from the axes emerges. The pattern is highly visible for 12 work piece holders and disappears in tests with 14 work piece holders and more. In the pattern, polygonal features are manually observable.

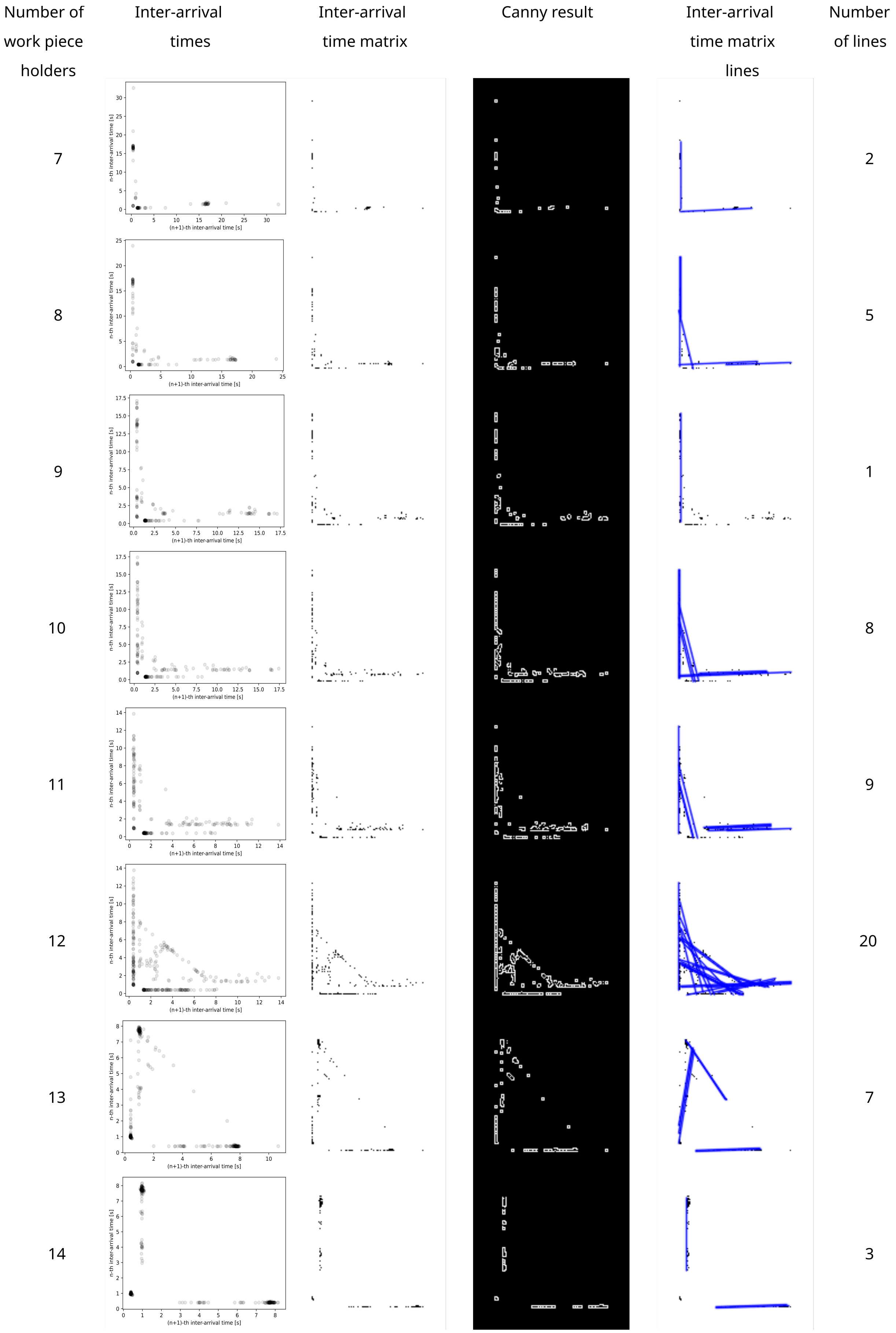


Figure 4: Test results after 15 min. The inter-arrival time matrix is derived from a histogram following Figure 2.

The lines that are detected by the method generally follow the visually observed patterns. However, more lines are generated than expected. In the case of 9 work piece holders, lines near the axes are not generated. Note that while line detection accuracy depends on the four parameters for the Probabilistic Hough Line Transform that are shown in Table 1, the most critical parameter has been the maximum gap between points. The result that the maximum number of generated lines clearly indicates the situation with the strongest pattern at 12 work piece holders is robust against parameter variations of less than ± 10% maximum gap between points.

6. Discussion

The proof of principle test demonstrates that for the considered scenario, the proposed method can be employed to automatically detect deterministic chaotic material flow behavior in double conveyor belts. The method and its algorithms that are state of technology in image analysis are able to extract lines that follow phase space trajectories.

However, while the number of detected lines grows with the convolution of the inter-arrival time pattern, the gaps in the pattern that remain after 15 min measurement time lead to misidentified lines. In the case of 13 work piece holders, lines are detected at the left side that go from bottom left to top right. Considering the other patterns, one would expect one line that is vertical and two parallel lines from top left to bottom right, which represent different routes through the system. These patterns become more dominant for longer test times. Effectively, the maximum gap between points parameters prevents correct detection of lines for the considered 15 min measurement time. However, reducing the maximum gap parameter results in a lower overall number of recognized lines to a point, in which the approach does not yield usable results. Therefore, the proposed approach is considered unsuitable for detecting root causes for deterministic chaotic behavior or for phase space trajectory reconstruction.

Nevertheless, automatic detection of chaotic behavior may provide insights for operative production management about system design shortcomings. The approach can easily be applied to existing conveyors e.g. by adding one proximity sensor per material flow loop. Employed at a large scale, it could draw attention to hidden issues that are normally covered by down times, breaks or idle times.

Besides its application using data from PLCs, the method can be directly applied to co-simulations e. g. in digital twins. Application in the design phase of systems would allow early checks for undesired deterministic chaotic behavior. As an addition to commonly used design quality checkers, the method could help to improve production system design.

7. Conclusion and Outlook

An approach for detecting deterministic chaotic inter-arrival times has been presented. In a test scenario, it has been shown to automatically detect patterns in inter-arrival time phase space. As a next step of research, the approach should be tested in real production situations that exhibit stochastic effects. If the results from the described laboratory test are reproducible in industrial production environments, the approach may help improve capacity flexibility by front loading issue handling. Instead of solving material flow bottlenecks that are caused by deterministic material flow chaotic behavior when production is maximized, detection may trigger a system or control redesign during normal operation.

Acknowledgements

The authors acknowledge the financial support from the European Regional Development Fund (EFRE) within the project SMAPS (grant number: 0200545).

References

[1] Leung, C., Lau, H., 2020. Simulation-based optimization for material handling systems in manufacturing and distribution industries. Wireless Networks 26.

[2] Ballarino, A., Brusaferri, A. et al., 2019. Knowledge Based Modules for Adaptive Distributed Control Systems, in: Tolio, T., Copani, G., Terkaj, W. (Eds.), Factories of the Future: The Italian Flagship Initiative. Springer International Publishing, Cham, pp. 83–108.

[3] Manns, M., ElMaraghy H.A., 2009. Inter-arrival time patterns in manufacturing systems with main and side loops. International Journal of Production Research 47 (10), 2721–2744.

[4] Dashkovskiy, S., Görges, M., Kosmykov, M., Mironchenko, A., Naujok, L., 2011. Modeling and stability analysis of autonomously controlled production networks. Logistics Research 3 (2), 145–157.

[5] Nyhuis, P. (Ed.), 2008. Beiträge zu einer Theorie der Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[6] ElMaraghy, H.A., Manns, M., 2007. Transition of interarrival time patterns between automated and manual configurations of assembly systems. Journal of Manufacturing Systems 26 (1), 1–11.

[7] Gunal, M.M., Karatas, M., 2019. Industry 4.0, Digitisation in Manufacturing, and Simulation: A Review of the Literature, in: Gunal, M.M. (Ed.), Simulation for Industry 4.0: Past, Present, and Future. Springer International Publishing, Cham, pp. 19–37.

[8] Lütjen, M., Rippel, D., 2015. GRAMOSA framework for graphical modelling and simulation-based analysis of complex production processes. The International Journal of Advanced Manufacturing Technology 81.

[9] Glatt, M., Aurich, J.C., 2019. Physical modeling of material flows in cyber-physical production systems. Procedia Manufacturing 28, 10–17.

[10] Zarco, L., Siegert, J., Schlegel, T., Bauernhansl, T., 2021. Scope and delimitation of game engine simulations for ultra-flexible production environments. Procedia CIRP 104, 792–797.

[11] Onvural, R.O., 1990. Survey of Closed Queueing Networks with Blocking. ACM Comput. Surv. 22 (2), 83–121.

[12] Hernández-González, S., Ramírez-Tapia, R., Jiménez-García, J.A., 2019. Analysis of the Productivity of a Shoe Production Line--Application of Queueing Theory and Lean Manufacturing, in: García Alcaraz, J.L., Rivera Cadavid, L. et al. (Eds.), Best Practices in Manufacturing Processes: Experiences from Latin America. Springer International Publishing, Cham, pp. 367–388.

[13] Ali A. Pouyan, Heydar Toossian Shandiz, Soheil Arastehfar, 2011. Synthesis a Petri net based control model for a FMS cell. Computers in Industry 62 (5), 501–508.

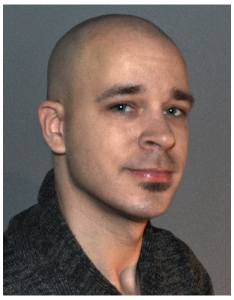
[14] Martínez-Olvera, C., Mora-Vargas, J., Campagnolo, A., 2018. A Max-Plus Algebra Approach to Study Time Disturbance Propagation within a Robustness Improvement Context. Mathematical Problems in Engineering 2018, 1–18.

[15] J. Matas, C. Galambos, J. Kittler, 2000. Robust Detection of Lines Using the Progressive Probabilistic Hough Transform. Computer Vision and Image Understanding 78 (1), 119–137.

[16] Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8 (6), 679–698.

Biography

Martin Manns (*1975) has been head of the Chair of Production Automation and Assembly (FAMS) at the University of Siegen since 2016. Before his appointment, Univ.-Prof. Dr.-Ing. Martin Manns has worked in production research at Daimler AG (2009-2016) and Henkel KGaA (2007-2009) and as a post doctorate fellow at the University of Winsor, Ontario, Canada (2006-2007).



Denny Höhnen (*1989) is currently studying for his master's degree in industrial engineering and is also working as a research assistant for the Chair of Production Automation and Assembly (FAMS) at the University of Siegen. Prior to that, he completed a training program to qualify as a state-certified technical engineer (2013-2015) and graduated with a bachelor's degree in industrial engineering (2015-2020).

3rd Conference on Production Systems and Logistics

Democratizing Manufacturing – Conceptualizing the Potential of Open Source Machine Tools as Drivers of Sustainable Industrial Development in Resource Constrained Contexts

Mohammed Omer[1], Melina Kaiser[2], Manuel Moritz[1], Sonja Buxbaum-Conradi[1], Tobias Redlich[1], Jens P. Wulfsberg[1]
[1]*Helmut-Schmidt-University, Hamburg, Germany*
[2]*University of Hamburg, Hamburg, Germany*

Abstract

Similar to open source software, the open source hardware (OSHW) movement is seen as a technology driver which can enable developing economies to leapfrog their industries. While machine tools are a subset of OSHW, they have received relatively little academic attention compared to electronic OSHW. This study applies an explorative research approach and analyses open source designs for machine tools freely available on the internet. By coining, the term open source machine tools (OSMT), it determines their applicability in low resource contexts and identifies the potential of OSMTs in democratizing manufacturing technologies. OSMTs thereby encourage diversification, entrepreneurship, and inclusive industrial development, thus contributing to the implementation of Sustainable Development Goal no. 8 which aims to promote inclusive and sustainable economic growth. Specific areas for OSMT application in low-resource contexts and factors and barriers affecting their success are singled out.

Keywords

Open Source Hardware; Democratizing Manufacturing; Open Source Machine Tools; SDG; Technology Driver; Resource Constrained Contexts; SME; Sustainability

1. Introduction

In the developing and least developed countries (LDCs), more than five billion people live in resource constrained contexts [1]. The global digital divide exacerbates this phenomenon: On the one hand, industrialized nations witness the evolution of industry 4.0, artificial intelligence, and associated technologies at a rapid pace; on the other hand, there is a persistent lack of technical know-how, digital infrastructure, and lack of scientific research capacities in developing countries and LDCs. As a consequence, this last group is left behind in global development. This has recently become even more evident as resource constrained environments have been hit hard by the disruption of globalized supply chains due to the COVID-19 pandemic [2]. People at the bottom of the socioeconomic spectrum, the so-called bottom of the pyramid, are affected most by this. These dynamics can be seen not only on the international but also on the national level.

To mitigate these inequalities, development theorists have frequently called for faster and more comprehensive inclusion of developing and LDCs into the global network of industrialized nations as well

DOI: https://doi.org/10.15488/12154

as for the creation of competitive environments to stimulate innovation and boost productive capacities[1] [3]. This is vital for sustainable and equitable economic growth that localizes production and explicitly includes communities and individuals at the bottom of the pyramid [2], especially as traditional developmental approaches such as dissemination of technology have often not yielded the intended successes.

Developments in the field of open source economics have recently offered opportunities to contribute to this goal. Not only are many software programs freely available for anyone to access, use, modify, distribute, and sell, but open source principles are increasingly applied to all kinds of tangible products in the form of open source hardware (OSHW). OSHW is a broad term that includes virtually any kind of tangible item that is produced based on information that is freely accessible, from toys and furniture to farming machinery. Research on OSHW has focused primarily on products such as medical devices [5,4], laboratory equipment [6,7], renewable energy systems [8,9] and also on improving education through OSHW [10,11]. However, even though it is precise about open source aspects, the OSHW definition makes no differentiations between the many different types of hardware that can be built based on open source principles. This paper advocates for the need of subclassifications catering to the particularities of the different kinds of products included in the OSHW definition and proposes the conceptualization of open-source machine tools (OSMT) as a distinct subdomain of OSHW. On the basis of this definition, this paper furthermore identifies OSMT's potential and challenges to democratize manufacturing and to contribute to the implementation of the SDGs and to enable emerging countries to leapfrog their economies through cost-effective, sustainable, and inclusive capacity building.

2. Background

2.1 The limits of traditional production technology

In industrial production technology practice, the innovation process has generally been closed source. As this innovation requires a high input of knowledge and capital, the process is largely barred to most SMEs in developing countries as they lack the necessary human and financial resources to carry out research and development [28]. Closed source production technology thus is a cohesive circle that hinders the dissemination of knowledge and capital to developing countries [12]. Many industrialized countries have in the last decades developed a consciousness for the intrinsic inequalities of traditional production technology and have put forward developmental efforts aimed at making advanced technologies available in developing contexts [13]. However, this top-down developmental aid has not always yielded the anticipated results [15,14]. Based on this observation, bottom-up economics utilizing the concept of co-creation enabled by participation and collaboration such as open source economics can be used as a means to achieve inclusive and sustainable development [16].

2.2 Open Source Hardware (OSHW)

With its roots in open source software (OSS), OSHW applies the open source principles to tangible objects such as machines, devices, or other physical artifacts [17]. The miniaturization of microcontrollers and electronics, coupled with increased accessibility to digital fabrication technologies lowering the barriers to rapid prototyping, have accelerated the evolution of OSHW technology. The democratization of the internet has allowed makers and creators to freely share their designs and know-how, while communities around the world replicate these designs, build upon them, and further improve them [18]. To qualify as OSHW, the design files required to reproduce the machine or device must be made openly available for anyone to study, modify, distribute, make, or sell [19]. Simple replication of an object only requires a bill of materials (BOM)

---

[1] Productive capacities are defined as "the productive resources, entrepreneurial capabilities and production linkages that together determine the capacity of a country to produce goods and services and enable it to grow and develop." [2].

and build instructions; editable documents such as computer-aided design (CAD) files in the native format are furthermore necessary to make modifications of OSHW projects. Other instructions such as assembly guidelines, drawings, or electrical schematics can also be made available. Based on this information, a user can either replicate a design, further develop it, or simply extract ideas for other projects [20]. In contrast to OSS, which relies exclusively on computers, OSHW requires tools, machines, and raw materials for replication. This increases OSHW's transaction costs and complexity compared to OSS [10].

In OSHW, two general domains can be distinguished, namely hedonic and utilitarian. With a majority of OSHW projects being hobbyist gadgets, DIY projects in the developed world can be classified as being predominantly hedonic in nature [21]. In contrast, in emerging economies, DIY machines mostly serve a utilitarian purpose [22]. Given that a majority of the OSHW projects are developed in industrialized countries and are designed according to their local resources, needs, and requirements, exact replications of these projects can be inefficient, uneconomical, or even unfeasible in the developing world [23]. Based on this observation, the concept of open source appropriate technology (OSAT) has emerged [24,18] which is based on the argument that advanced technologies are not appropriate for the contexts of developing nations. Instead of exporting the technologies of industrialized countries, technologies should be developed to fit the specific cultural, environmental, socioeconomic, and educational contexts in which they are to be applied [25]. Appropriate technology is characterized by having low capital demands, being locally controlled, decentralized, small-scale, labour-intensive, and energy-efficient [26]. OSAT's goals to adapt technology to human needs and specific socioeconomic and cultural contexts are nowadays also often seen in relation to sustainable and environmentally sound technologies [27].

3. Building on the Shoulders of Giants: Conceptualizing Open Source Machine Tools (OSMT)

One category of items that has in recent years increasingly been built based on open source principles are machine tools. The term machine tools has varying definitions but has generally referred to forming, milling, or grinding machines with a focus on metal processing. Machine tools are directly or indirectly used to make every modern human-made object and with their ability to produce the components required to make other machines, machine tools are aptly known as "mother machines" [28]. Microelectronics and computer technology have allowed the digital control of machine tools, introducing the nowadays widely used Computer Numerical Control (CNC) systems. "Digital fabrication machines" is another umbrella term that includes all machines that can be digitally controlled, whereby laser cutters and 3D printers are also included along with conventional CNC machine tools [27]. For the sake of conciseness and to allow for a certain degree of generalization, we will refer to all these machines – CNC controlled or not, hobbyist or industrial grade – that enable the production of products simply as machine tools.

With their importance for different industries and their potential to bridge sectors, technological developments in machine tools are regarded to generally have the highest impact on the productivity of economic systems compared to innovations in other fields [29,30]. However, only a few industrialized nations are the sole producers of machine tools and manufacturing technologies [31]. Many developing countries lack the resources and know-how to invent and produce such technologies themselves thus rely on buying them abroad. This, however, requires not only high upfront capital investments, but these machine tools are often difficult and expensive to import into low-income countries due to high shipping costs, customs taxes, and administrative hurdles. A lack of technical know-how and low technological literacy[2] further exacerbate these challenges. Taken together, the cost of purchasing and using a high-tech machine tool is prohibitively high in resource constrained contexts; this confirms the assumptions of appropriate

---

[2] The ability to work with technology in different ways, from understanding over accessing and using to managing it, is called technological literacy or technology literacy. For an individual to obtain technological literacy, a certain degree of digital literacy is an important precondition. Via digital devices such as computers, said individual is enabled to make use of digital information in various forms that is accessible on the internet [32].

technology. Therefore, even though machine tools can be regarded as decisive technology drivers for economic growth that have the potential to leapfrog developing economies, high-tech machine tools are often unattainable in developing contexts.

The OSHW movement offers a potential solution to this problem: By applying OSHW principles to machine tools and making build instructions, BOMs, electronic schematics, and CAD files for machine tools freely available on the internet, the divide in their accessibility between industrialized and developing economies could be bridged and bottom-up, community-driven developmental approaches could be strengthened. While open source machine tools already exist in myriad forms, they have, however, not previously been defined as such. The concept of open source machine tools was first used by Pat Delany, who designed and developed OSHW projects focused on machine tools for "do-it-yourself global development" [33]. Recent years have also seen collective efforts such as Open Source Ecology's Global Village Construction Set [34], which is a set of 50 open source industrial machines and machine tools required to build a small but sustainable civilization, as well as individual efforts from the maker community such as the PrintNC open source CNC mill [35] and the Voron, an open source 3D printer [36]. By applying OSAT design principles and open collaboration, these projects have been able to develop machine tools for a fraction of the cost of their commercial counterparts, and they have been successfully replicated by the community [37]. Online repositories like the Open Hardware Observatory have compiled a list of self-built machine tools collected from the internet by means of a crawlerbot[3] using keywords such as 'DIY' and 'homemade', and have sorted them into various machine tool categories [39]. These DIY or homemade machine tools range from desktop machines weighing a few kilograms to large scale industrial machines. Out of the thousands of designs online, only a handful (10 – 20 machine tools per category) fulfil the OSHW requirements as most do not publish any blueprints to allow replication [40].

However, despite open-source machine tools' great potential for inclusive economic growth and their popularity among the maker community, they have hitherto been neglected in academic OSHW research. This paper therefore proposes the introduction of open-source machine tools (OSMT) as a distinct concept under the domain of OSHW to reflect their importance, to encourage research and standardization efforts in the field and to foster dissemination. To meet the minimal requirements to qualify as OSMT, an item must be a machine tool as defined above. Parallel to the OSHW definition, a machine tool can be considered to be open source if all the plans and instructions that are needed to reproduce it are publicly available under free terms and for anyone. Furthermore, the project source files should be easily accessible to allow for community collaboration in the development, replication, and feedback processes. Due to the open development process of these products and the holistic and dynamic involvement of stakeholders in all stages from ideation over production to modification, OSMT can be considered a disruptive innovation.

Considering the complexity of machine tools, the design sharing function is a key differentiator of OSMT from proprietary machine tools, whereby end users can save tremendous research and development efforts by building on the experience or findings of other users, thus avoiding constant reinventions [41]. Since OSMT are built by the end-users, they are then also able to maintain and fix any issues that could arise during the lifetime of the machines. Production occasionally requires custom machine tools for which there might not be an ideal solution on the market and that is when a self-built machine tool can optimally use available workshop area and be built according to user's specific needs and requirements. With most open source practitioners being individuals working from their homes or garages, the parts and resources for these machines are often cheaper options compared to their industrial counterparts, which can sometimes cause reliability issues. Moreover, as many of the designers are no engineers, there is also often a lack of engineering best practices, which can lead to either under- or overengineered solutions. However, through standardization efforts and focused research in these fields, in the future, OSMT have the potential to attain

---

[3] Webcrawling is the technical term for automatically accessing a website and obtaining data via a software program [38].

the reliability, performance, and quality normally associated with commercial machine tools. Projects like the Helmut Schmidt University's Open Lab Starter Kit, which aims at designing meticulously documented and globally replicable open source machine tools that comply with industry norms hence prioritizing user safety, can help this process.

4. OSMTs' role in implementing the SDGs

OSHW can constitute a sustainable developmental alternative for developing industries by fulfilling the criteria of OSAT which combines the principles of appropriate technology with open source characteristics. Especially the subdomain of OSMT has great potential for this as it is able to "meet the boundary conditions set by environmental, cultural, economic, and educational resource constraints of the local community" [18]. OSMT developed under OSAT principles – by making use of locally available resources, using off the shelf components, offering customization towards specific user needs as well as expandability or upgradability, integrating community support to answer technical questions – are affordable, easily accessible, and flexible. This democratization of machine tool access can lead to more inclusive and sustainable economic growth by offering opportunities for new business models. It is therefore directly related to the implementation of the Sustainable Development Goals (SDGs) [42]. Moreover, OSMT can be the key to building local production capacity which is vital to realize the goals of achieving a circular economy by localizing production [43].

OSMT can be regarded as a key enabling technology that lowers the barriers to manufacturing in developing countries. By opening up opportunities to prototype, design, create, and invent products that are intimately adapted to local markets and needs, local entrepreneurship is strengthened. As machine tools, that would normally have been imported, can be produced locally based on OSHW principles, local economies become more resilient and less vulnerable to economic shocks through reduced import dependency. It eliminates the need for high upfront capital investments and facilitates a quick return on investments as OSMT are cheaper compared to conventionally produced machine tools [37]. This further encourages economic diversification by offering more business opportunities with less risk, especially for SMEs with little human and capital resources, all the while strengthening resilience by offering options for circular economic models. In this, OSMT contribute to the implementation of SDG no. 8, and especially target 8.2, which aims at promoting "sustained, inclusive, and sustainable economic growth […]" by achieving "higher levels of economic productivity through diversification, technological upgrading and innovation" [44].

Next to bridging differences in the economic development on an international level, the adoption of OSMT also has the potential to bridge national inequalities. In this, it helps to achieve SDG no. 10, which targets the reduction of "inequality within and among countries", by contributing to sub-target 10.2. ("[…] empower and promote the social, economic and political inclusion of all, irrespective of age, sex, disability, race, ethnicity, origin, religion or economic or other status") [44]. Xenophobic national economic policies can disadvantage certain groups of the society and exclude them from equitable economic development [45]. Here, OSMT can empower marginalized communities that are failed by conventional development policies to become self-reliant and pursue bottom-up economic growth, thus making inclusive and fair development a reality. To realize this potential of OSMT and to use them to address the SDG implementation, there is an additional need for new educational models and incentives.

5. Potential Challenges to OSMT Adoption

While further research is necessary to determine the barriers to OSMT adoption in specific local, resource-constrained contexts, some general challenges can already be anticipated. One basic factor hindering the adoption of open-source technology is the lack of awareness of OSHW in general and particularly OSMT. Limited access to the internet and low digital literacy decrease accessibility as OSHW projects are often strewn over the internet across various forums, social media platforms and online repositories [46]. To meet this challenge, it has been suggested that a global centralized open source database be created to house all OSHW projects that are relevant to the UN Sustainable Development Goals (SDGs) [46].

A further challenge is the replicability of OSMT projects. Encompassing the fields of mechanics, electronics, and software, building a machine tool can be a complex endeavour requiring proficiency in various domains such as precision fabrication, high voltage electrical wiring and software configuration. In contrast to OSS, where replication mostly involves downloading and running a code, the replicability of OSMT depends on more pre-requisites such as required manufacturing tools, accessibility of parts and raw materials, as well as considerably more technical know-how, digital literacy, money, and time. An OSMT project can be considered replicable if a functional version of the project can be built by builders in their respective locations, which is also a key characteristic of the OSHW definition [47]. As user requirements and access to parts, raw materials, and manufacturing capability differ in different regions of the world, developing globally accessible designs that are also locally adaptable remains a challenge. Language barriers and cultural differences are further factors that potentially hinder replicability.

OSMT need to moreover be designed in a way that people with little formal education are also able to access, manufacture, assemble, and use the machine tools. The local unavailability of building materials and machine tools to manufacture components in the context of limited resources could potentially hinder the replication of a product that could easily be built in non-constrained circumstances. This is especially likely for electronic components and precision machined components. The arising need to import the building materials from elsewhere counteracts many of the advantages of OSMT described earlier. However, the increase in locally built OSMT could potentially create a market pull for local suppliers of machine tool components to emerge, which creates the possibility of new supplier networks and corresponding business models.

A further challenge lies in the fact that many projects are also outdated, provide insufficient instructions, or lack robustness and quality, thus hindering replication. Without meticulous documentation of the fabrication methods and assembly guidelines, the machine tool might turn out imprecise or unusable or the process could pose hazards to the safety of the builder. One aspect that is specific to the design of proprietary machine tools for industrial use is that the machines are designed with design standards and norms such as the CE certification that ensures compliance with the EU machinery directive (2006/42/EC) and hence the safety of the end user. With no regulatory requirements to abide by, the user safety aspect is commonly neglected in the development of OSMT, since they are often designed for personal use in provisional workshops and therefore do not need to conform to any industry norms. Without the need to commercialize the machine, there is often no incentive to invest resources, time, and money on the extra development effort required to build industry-compliant machines. Organizations such as the open hardware repository have developed a community-based OSHW certification process which is carried out by experts who check designs for user safety and reliability, thus aiming to develop quality standards for OSHW [48]. A similar process to verify the safety and reliability of OSMT is paramount to safe diffusion of these technologies, but the complexity of the machine tools is likely to make this a challenge. Moreover, new business models need to be derived that gives OSMT practitioners sufficient incentive to develop well-documented OSMT that also take user safety and industry standards into consideration.

6. Conclusion and Future Works

Coining the concept of OSMT shows the potential and necessity for subclassifications under the domain of OSHW. This is hoped to stimulate future research specifically on machine tools' particularities in open source contexts as well as encourage further conceptualizations of other types of OSHW. By defining OSMT as a distinct concept, it is also possible to analyse their specific potential for sustainable, context-appropriate, and inclusive development and their ability to contribute to the implementation of the SDGs.

As disruptive innovations, OSMT are here not intended as a competition to industrial machine tools whose technological efficiency and precision have evolved over decades, but rather as a means of making machine tools more accessible and affordable to communities and individuals who would otherwise not be able to use machine tools at all, thus democratizing manufacturing. This, in turn, can strengthen the local job market, re-shore manufacturing, lower import dependence, and create more resilient, equitable, and inclusive economies in emerging contexts. By doing so, OSMT allow developing, resource constrained economies to leapfrog their industries in a sustainable and inclusive way, making the innovations of Industry 4.0 accessible to them, and helping to close the ever-widening technological gap.

However, the successful adoption of OSMT requires more than just engineering innovation. Technology adoption and diffusion in developing economies are complex socio-technical issues which demand a multi-disciplinary approach that incorporates the diverse perspectives of engineering, social sciences, developmental economics, and governance. Such holistic approaches, in which the roles of various stakeholders, public values, and technological innovations are taken into account [49], are important because, even with the right intentions, innovation applied wrong can potentially lead to social exclusions and increased inequalities [50,51]. While being wary of falling into the traps of paternalism [4], policy-driven top down approaches as well as community-aligned bottom up approaches need to be considered to truly facilitate inclusive capacity building in emerging economies [16]. To further evaluate the potential of OSMT and determine barriers and challenges associated with their adoption, pilot projects and associated research such as the Open Lab Starter Kit [40] will need to be carried out and monitored in the long term.

Acknowledgements

The authors would like to thank the Center for Digitalization and Technology Research of the Bundeswehr (DTEC.bw) for their support.

References

[1] Collier, P., 2007. The bottom billion: Why the poorest countries are failing and what can be done about it. Oxford University Press.

[2] UNCTAD, 2021. Building productive capacities critical for least developed countries. https://unctad.org/news/building-productive-capacities-critical-least-developed-countries. Accessed 4 January 2022.

[3] Lundvall, B.-Å., Vang, J., Joseph, K.J., Chaminade, C., 2009. Bridging innovation system research and development studies: challenges and research opportunities. 7th Globelics Conference.

[4] Niezen, G., Eslambolchilar, P., Thimbleby, H., 2016. Open-source hardware for medical devices. BMJ innovations 2 (2), 78–83.

[5] Moritz, M., Redlich, T., Günyar, S., Winter, L., Wulfsberg, J.P., 2019. On the Economic Value of Open Source Hardware – Case Study of an Open Source Magnetic Resonance Imaging Scanner. Journal of Open Hardware 3 (1).

[6] Fisher, D.K., Gould, P.J., 2012. Open-source hardware is a low-cost alternative for scientific instrumentation and research. MI 1 (2), 8–20.

[7] Pearce, J.M., 2012. Building research equipment with free, open-source hardware. Science (New York, N.Y.) 337 (6100), 1303–1304.

[8] Carballo, J.A., Bonilla, J., Roca, L., Berenguel, M., 2018. New low-cost solar tracking system based on open source hardware for educational purposes. Solar Energy 174, 826–836.

[9] Gazzo, A., Gousseland, P., Verdier, J., Kost, C., Morin, G., Engelken, M., Schrof, J., Nitz, P., Selt, J., Platzer, W., Ragwitz, M., Boie, I., Hauptstock, D., Eichhammer, W., 2011. Middle East and North Africa region assessment of the local manufacturing potential for Concentrated Solar Power (CSP) projects, 223 pp. Accessed 26 November 2021.

[10] Arancio, J.C., 2020. Opening up the tools for doing science: The case of the global open science hardware movement. Preprint. https://osf.io/46keb. Accessed 22 November 2021.

[11] Kostakis, V., Niaros, V., Giotitsas, C., 2015. Open source 3D printing as a means of learning: An educational experiment in two high schools in Greece. Telematics and Informatics 32 (1), 118–128.

[12] Boldrin, M., Levine, D.K., 2001. Against intellectual monopoly. Cambridge University Press, Cambridge.

[13] Wilson, G., 2007. Knowledge, innovation and re-inventing technical assistance for development. Progress in Development Studies 7 (3), 183–199.

[14] Salazar Xirinachs, J.M. (Ed.), 2014. Transforming economies: Making industrial policy work for growth, jobs and development. International Labour Office, Geneva, 402 pp.

[15] Banerjee, A.V., Duflo, E., 2011. Poor economics: A radical rethinking of the way to fight global poverty / Abhijit V. Banerjee and Esther Duflo. PublicAffairs, New York.

[16] Redlich, T., Moritz, M., Wulfsberg, J.P. (Eds.), 2019. Co-creation: Reshaping business and society in the era of bottom-up economics. Springer International Publishing, Cham.

[17] Balka, K., 2011. Open source product development: The meaning and relevance of openness. Zugl.: Hamburg-Harburg, Techn. Univ., Institut für Technologie- und Innovationsmanagement, Diss., 2011, 1st ed. Gabler, Wiesbaden, 196 pp.

[18] Pearce, J.M., 2012. The case for open source appropriate technology. Environment, Development and Sustainability 14, 425–431.

[19] Open Source Hardware Association, 2020. Open-source hardware FAQ. https://www.oshwa.org/faq/. Accessed 6 January 2022.

[20] Tseng, T., Resnick, M., 2014. Product versus process, in: DIS'14. Proceedings of the 2014 ACM SIGCHI conference on designing interactive systems. DIS '14: Designing Interactive Systems Conference 2014, Vancouver BC Canada. 21 06 2014 25 06 2014. ACM, New York, NY, pp. 425–428.

[21] Hansen, A., Howard, T.J., 2012. The Current State of Open Source Hardware: The Need for an Open Source Development Platform, in: Chakrabarti, A., Prakash, R.V. (Eds.), Icord'13. Springer, New York, pp. 977–988.

[22] Powell, A., 2012. Democratizing production through open source knowledge: from open software to open hardware. Media, Culture & Society 34 (6), 691–708.

[23] Tanenbaum, T.J., Williams, A.M., Desjardins, A., Tanenbaum, K. Democratizing technology: Pleasure, utility and expressiveness in DIY and maker practice, in: CHI 2013: Changing Perspectives, Paris, France.

[24] Hazeltine, B., Bull, C., 2003. Field guide to appropriate technology. Academic Press.

[25] Schumacher, E.F., 1973. Small is beautiful: A study of economics as if people mattered. Blond & Briggs, London.

[26] Akubue, A., 2000. Appropriate technology for socioeconomic development in third world countries. JOTS 26 (1).

[27] DuBose, J., Frost, J.D., A Chameau, J.-L., A Vanegas, J., 1995. Sustainable development and technology, in: The environmentally educated engineer. Focus on fundamentals. Workshop on the fundamentals of environmental engineering education, Christchurch, NZ. 22-24.08.1994.

[28] Mori, M., Hansel, A., Fujishima, M., 2014. Machine tool, in: Laperrière, L., Reinhart, G. (Eds.), CIRP Encyclopedia of Production Engineering. Springer, pp. 792–801.

[29] Rosenberg, N., 1983. Inside the black box: Technology and economics. Cambridge University Press, Cambridge.

[30] Saxena, P.K., Sharma, A., 2014. Role of machine tool industry in economic development. ISSN 3 (5), 188–193.

[31] Békés, G., Harasztosi, P., 2020. Machine imports, technology adoption, and local spillovers. Rev World Econ 156 (2), 343–375.

[32] Technology for All Americans Project, International Technology Education Association, 2002. Standards for technological literacy: Content for the study of technology, 2nd ed. International Technology Education Association, Reston, Va.

[33] Delany, P. Do-it-yourself global development. Open Source Machine Tools. http://opensourcemachinetools.org/wordpress/open-source-tools-diy-global-development/. Accessed 14 January 2022.

[34] Moritz, M., Redlich, T., Grames, P.P., Wulfsberg, J.P., 2016 - 2016. Value creation in open-source hardware communities: Case study of Open Source Ecology, in: . 2016 Portland International Conference on Management of Engineering and Technology (PICMET), Honolulu, HI, USA. 9/4/2016 - 9/8/2016. IEEE, pp. 2368–2375.

[35] PrintNC Wiki. PrintNC V3.0. https://wiki.printnc.info/en/home. Accessed 28 January 2022.

[36] Maks Zolin, Tin Pecirep and VORON Design contributors, 2022. VORON Design. https://vorondesign.com/. Accessed 2 January 2022.

[37] Open Source Ecology. Machines: Global village construction set. https://www.opensourceecology.org/gvcs/. Accessed 9 January 2022.

[38] Cloudfare. What is a web crawler? How web spiders work. https://www.cloudflare.com/learning/bots/what-is-a-web-crawler/. Accessed 1 February 2022.

[39] Open Hardware Observatory, 2022. Search engine for sustainable open hardware projects. https://en.oho.wiki/wiki/Category:Business,_industry. Accessed 2 January 2022.

[40] Omer, M., 2021. Open Lab Starter Kit: A OSHW Repository with blueprints and plans for low marginal cost replication of Open Source Machine Tools (OSMT). Helmut Schmidt University. https://hardware.development.fabcity.hamburg/open-lab-starter-kit/. Accessed 2 March 2022.

[41] Lowe, A.S., 2019. Distributed manufacturing: Make things where you need them, in: Redlich, T., Moritz, M., Wulfsberg, J.P. (Eds.), Co-creation. Reshaping business and society in the era of bottom-up economics. Springer International Publishing, Cham, pp. 37–50.

[42] United Nations Economic Council. Resolution 2021/30: Open-source technologies for sustainable development, 2 pp.

[43] Joshua Pearce, 2021. An Open Source Preemptive Strike in the Coming War Over The Freedom to Make Your Own Products. Cosmolocal Reader.

[44] United Nations, 2021. The 17 goals. United Nations. https://sdgs.un.org/goals. Accessed 8 December 2021.

[45] Jureidini, R., 2005. Migrant workers and xenophobia in the Middle East, in: Bangura, Y., Stavenhagen, R. (Eds.), Racism and public policy. Palgrave Macmillan UK, London, pp. 48–71.

[46] UNCTAD (Ed.), 2021. Note on a proposed UN centralised database of open-source appropriate technologies. Produced for the development of Resolution E/RES/2021/30 entitled "Open-source technologies for sustainable development" adopted by the United Nations Economic and Social Council on 22 July 2021, 32 pp.

[47] Antoniou, R., Pinquié, R., Boujut, J.-F., Ezoji, A., Dekoninck, E., 2021. Identifiying the factors affecting the replicability of open source hardware designs, in: Proceedings of the International Conference on Engineering Design (ICED21), Gothenburg, Sweden. 2021, pp. 1817–1826.

[48] Open Hardware Observatory, 2020. About OHO Open Hardware Observatory. Open Hardware Observatory. https://en.oho.wiki/wiki/About_OHO_Open_Hardware_Observatory. Accessed 23 January 2022.

[49] Redlich, T., Buxbaum-Conradi, S., Basmer-Birkenfeld, S.-V., Moritz, M., Krenz, P., Osunyomi, B.D., Wulfsberg, J.P., Heubischl, S., 2016 - 2016. OpenLabs: Open source microfactories enhancing the FabLab idea, in: 2016 49th Hawaii International Conference on System Sciences (HICSS). 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA. 1/5/2016 - 1/8/2016. IEEE, pp. 707–715.

[50] Arocena, R., Senker, P., 2003. Technology, Inequality, and Underdevelopment: The Case of Latin America. Science, Technology, & Human Values 28 (1), 15–33.

[51] Cozzens, S.E., Thakur, D. (Eds.), 2014. Innovation and inequality: Emerging technologies in an unequal world. Edward Elgar, Cheltenham.

Biography

Mohammed Omer, M.Sc., studied mechanical engineering at the Technical University of Dortmund with a focus on manufacturing technologies. He is a doctoral candidate at the Institute of Production Engineering (LaFT) at Helmut-Schmidt-University in Hamburg. His research focuses on open source hardware development in the field of machine tools.

Melina Kaiser, M.A., has studied Languages and Cultures of Southeast Asia at the University of Hamburg and pursues a PhD in the same field. Her research focuses on environmental justice, equitable resource use, and sustainable development in the Vietnamese context.

Dipl-Ing. **Manuel Moritz**, MBA, is a research assistant at the Institute of Production Engineering at Helmut-Schmidt University, Hamburg and member of the working group value creation systematics. His research focus is value co-creation in online communities with or without organizations involved.

Dr. rer. nat. **Sonja Buxbaum-Conradi**, is a senior researcher at the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in Hamburg since 2013. Her research focusses on Global production networks, Knowledge processes in distributed value creation networks as well as Open production and open source hardware communities.

Dr.-Ing. Dipl.-Wirt.-Ing. **Tobias Redlich**, MBA is the chief engineer at the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in Hamburg since 2013. His research focuses on value creation systematics and the impact of new value creation patterns on production engineering and production management

Univ.-Prof. Dr.-Ing. **Jens P. Wulfsberg** is head of the Chair of Production Engineering in the Department of Mechanical Engineering at Helmut-Schmidt University in Hamburg since 2001. The main research areas at the Manufacturing Technology Laboratory (LaFT) are value creation systems, robotics, micro manufacturing as well as additive manufacturing and lightweight construction

3rd Conference on Production Systems and Logistics

Pareto Heuristic For Product Family-Oriented Product-Workstation Allocation Planning With Restricted Capacities

Christian Urnauer[1], Anna-Maria Larem[1], Joachim Metternich[1]
[1]*Institute of Production Management, Technology and Machine Tools, Darmstadt, Germany*

Abstract

A prerequisite for value stream design is the segmentation of products into families. This means that all products of a family are assigned to the same group of workstations so that the resulting material flows are separated as good as possible, and a higher degree of transparency is reached on the shop floor. However, since the number of workstations as well as their capacity is limited, shared resources cannot always be avoided in practice. Furthermore, the objective of product family orientation may compete with the objective of fulfilling product-workstation preferences. These preferences result, for example, from required equipment like specific tooling or from capability requirements. An optimization heuristic for this product-workstation allocation problem is presented within this article. First, the mathematical problem is formally described, then the heuristic is introduced and the required data for its application is outlined. For the evaluation, an extensive test set is generated, comparison heuristics are implemented, and solutions are made comparable through problem-specific bounds for both objectives. The results show that the solution quality of the pareto heuristic for both objective functions achieves almost the level of the comparison heuristics, which optimize only one of the objectives in isolation.

Keywords

value stream design; segmentation; pareto optimization

1. Introduction

Value stream design (VSD) is a widely used method to re-design material and information flows with the target of achieving a leaner, waste-reduced future state. The method is characterized by its simplicity and its application has led to significant improvements throughout different industries [1]. Practitioners particularly value the transparency and improvements achieved by the application of VSD [2]. However, in complex production environments, such as in high variety - low volume type companies, the application of the standard method is challenging [3]. Therefore, the increasing data availability in production can be facilitated to holistically support the value stream method with data analytics [4].

The segmentation of product families is a prerequisite for VSD. Families are identified by the similarity of the products' process steps in the downstream segment of the value stream, usually supported by a product-process-matrix [5]. In the segmentation, resources of each process are assigned to the product families with regards to their demand. This way, material flows are unbundled, the shop floor is structured, clear responsibilities can be assigned, departmental interfaces are reduced, and simpler planning and control mechanisms as well as the easier identification of improvements are facilitated through a higher degree of transparency [6]. However, experience from various VSD projects shows, that this step is frequently neglected and there are two major reasons for that:

DOI: https://doi.org/10.15488/12193

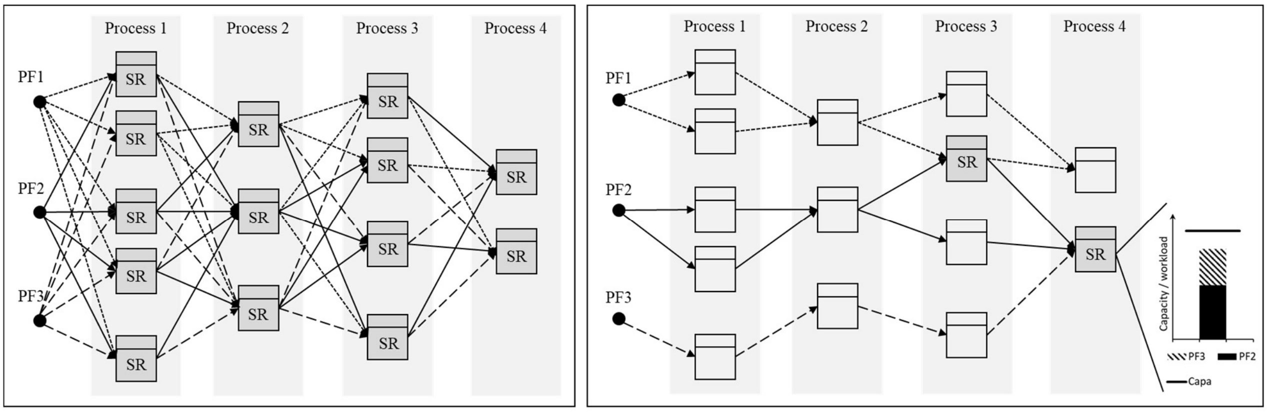
Firstly, identifying product families is particularly challenging if the product portfolio is very big and material flows are complex. This challenge can be countered with data analytics approaches applied to the product-process matrix [7]. Since the matrix can contain hundreds or thousands of products in mixed model value streams, a computer-aided sorting logic can help to visualize product family affiliations [8]. Furthermore, the application of cluster algorithms supports the identification of product families [4,7,9].

The second challenge in value stream segmentation is to realize a good product-workstation allocation that results in separable segments for the determined product families while considering workstation preferences and capacity restrictions. Following the standard design approach, product demand and workstation capacities are initially not taken into account and therefore, solutions are not practicable and must be adjusted iteratively [6]. Smaller adjustments can be achieved through optimizations in processing times, setup times, or availability. For bigger adjustments, operating times, shift models, the number of workers or machines must be increased or decreased, or products must be moved to other workstations with overcapacity. Depending on the initial situation, this typically results in a value stream design with shared resources and unevenly levelled workload. Additionally, workstation prioritizations are not taken into account in the standard approach. These prioritizations may result from restrictions such as required equipment, available workspace, degree of automation, or simply from personal preferences. In order to develop value stream oriented design proposals for such multi-dimensional solution spaces in complex production environments, smart optimization heuristics are required [10].

To support a product family-oriented segmentation under consideration of the available capacities, this paper proposes a heuristic planning approach. The multi-objective optimization problem as well as its restrictions are formulated in the next section, followed by the description of the product-workstation allocation planning heuristic and an outline of the required business data. Eventually, the approach is evaluated with a generated test set using assessment measures to estimate the solution quality in relation to the solution space and comparison heuristics that allow for contrasting with other feasible solutions.

2. Problem Definition

As illustrated in Figure 1, the product family-oriented segmentation transfers unclear material flows (left) to a clearly structured shop floor with fewer shared resources (right). This target is reached through a well-planned assignment of products to workstations which are limited in capacity. In the lower right of the figure, it is illustrated how the workload of two product families adds up on a shared resource but is covered by the slightly higher capacity of the respective workstation. Another aspect to consider in a well-planned product-workstation allocation is the workstation prioritizations that may exist for the products to be allocated.



PF = product family; SR = shared resource

Figure 1: Segmentation of the material flows of three product families

An ideal solution is characterized by:

A) an unambiguous *dedication of workstations to product families* so that clearly separatable segments are created with as few shared resources as possible, where
B) as many products as possible are *assigned to their most prioritized workstation.*

With the two independent objectives *product family separation* and *prioritization-based assignment*, this is a pareto optimization problem. It is characteristic for these kinds of problems that there is a set of pareto optimal solutions - the so called pareto front, for which no objective value can be improved without worsening the other. An optimal choice among these solutions depends on the decision maker's preferences.

The restrictions to this allocation problem are:

- every product belongs to one product family
- every product must be assigned to one workstation for every process it passes
- the total load of all assigned products must be covered by the capacity of the workstation

The mathematical problem formulation is outlined in the following. The objectives are described in section 0 and the restrictions are discussed in section 2.2. Since the product-workstation allocation for each process in the value stream is independent of the other processes' allocation, the problem is formulated for the consideration of one process. The used notations are summarized in Table 1.

Table 1: Notations

$FWA_{f,ws}$	Binary variable that indicates whether at least one product of family f is allocated to workstation ws
FWA_{total}	Total number of assignments of product families f to workstations ws
$Prio_{p,ws}$	Prioritization value for product p to be allocated to workstation ws
$Prio_{total}$	Total prioritization reached by all product-workstation assignments
$\emptyset\ Prio_{f,ws}$	Average prioritization value of all unassigned products of a family f for a workstation ws
$x_{p,ws}$	Binary decision variable that indicates whether product p is allocated to workstation ws
$z_{p,f}$	Binary parameter that indicates whether product p is part of product family f
y_p	Binary parameter that indicates whether product p is produced on this process
L_{ws}	Workload of workstation ws
C_{ws}	Capacity of workstation ws
$ct_{p,ws}$	Cycle time of product p at workstation ws
d_p	Demand per week for product p at the considered process
s_{ws}	Number of shifts for workstation ws
ot_{ws}	Operating time of workstation ws
OEE_{ws}	Overall equipment effectiveness for workstation ws
F	Set of product families
P	Set of products
WS	Set of workstations
f	Index for product families
p	Index for products
ws	Index for workstations

2.1 Objectives

$$Minimize\ FWA_{total} = \sum_{ws\in WS} \sum_{f\in F} FWA_{f,ws} \tag{1}$$

$$Maximize\ Prio_{total} = \sum_{ws\in WS} \sum_{p\in P} Prio_{p,ws} \cdot x_{p,ws} \tag{2}$$

with:

$$FWA_{f,ws} = \begin{cases} 1, & if \sum_{p \in P} x_{p,ws} \cdot z_{p,f} > 0 \\ 0, & else \end{cases} \qquad \forall\, ws \in WS, \forall\, f \in F \qquad (3)$$

The binary decision variable of the problem is $x_{p,ws}$, which indicates whether product p is assigned to workstations ws ($x_{p,ws}$= 1) or not ($x_{p,ws}$= 0).

Equation (1) expresses objective A): the number of assignments of product families f to workstations ws must be minimized, so that resources are shared with as few product families as possible. These family-workstation assignments are described by the binary variable $FWA_{f,ws}$. In equation (3), $FWA_{f,ws}$ takes the value 1, if at least one item of product family f is assigned to workstation ws, and 0 otherwise. The binary parameter $z_{p,f}$ in this equation indicates whether a product p belongs to the product family f. Thus, the product of $x_{p,ws}$ and $z_{p,f}$ is 1 if a product of the considered family is assigned to the workstation ws via $x_{p,ws}$. Summing $FWA_{f,ws}$ for the set of all existing families F in equation (1) gives the number of families sharing a workstation ws. Summing this number of planned product families per workstation for the set of all workstations WS, results in the total number of family-workstation assignments to be minimized.

Equation (2) expresses objective B): assignments must fulfil the highest possible product-workstation prioritizations. $Prio_{p,ws}$ is a parameter indicating the preferences for each product p to be produced on workstation ws. In practice, these prioritizations can be collected in the form of a product-workstation prioritization matrix [7], in which $Prio_{p,ws}$ can take any real number. However, it makes sense to use a predefined scale for the value range when collecting the prioritizations. The product of $Prio_{p,ws}$ and $x_{p,ws}$ is summed for all products and workstations, so the result is a total prioritization value to be maximized.

2.2 Restrictions

$$\sum_{f \in F} z_{p,f} = 1 \qquad \forall\, p \in P \qquad (4)$$

$$z_{p,f} \in [0,1] \qquad \forall\, p \in P, \forall\, f \in F \qquad (5)$$

$$\sum_{ws \in WS} x_{p,ws} = y_p \qquad \forall\, p \in P \qquad (6)$$

$$x_{p,ws} \in [0,1] \qquad \forall\, p \in P, \forall\, ws \in WS \qquad (7)$$

$$y_p \in [0,1] \qquad \forall\, p \in P \qquad (8)$$

$$L_{ws} \leq C_{ws} \qquad \forall\, ws \in WS \qquad (9)$$

with:

$$L_{ws} = \sum_{p \in P} x_{p,ws} \cdot ct_{p,ws} \cdot d_p \qquad \forall\, ws \in WS \qquad (10)$$

$$C_{ws} = s_{ws} \cdot ot_{ws} \cdot OEE_{ws} \qquad \forall\, ws \in WS \qquad (11)$$

Equations 5, 7, and 8 define $z_{p,f}$, $x_{p,ws}$ and y_p to be binary. Since one product can only belong to one product family, the sum of $z_{p,f}$ over the set of all families F must be 1 for each product p, which is expressed by equation 4. If a product p passes the considered process during its production, the parameter y_p is 1.

Therefore, the sum of product-workstation assignments $x_{p,ws}$ over all workstations WS of this process must be equal to y_p (equation 6). This ensures that one workstation is assigned for each required process. Equation 9 is the capacity restriction. The total load of a workstation L_{ws} must always be less than or equal to its available capacity C_{ws}. The workload of each product is calculated as its demand at the considered process d_p multiplied with its cycle time at the considered workstation $ct_{p,ws}$. Equation 10 constructs the total load as the sum of all product assignments $x_{p,ws}$ multiplied with their cycle times and demands. Eventually, equation 11 gives the total capacity of a workstation as the product of its number of shifts s_{ws}, its available operating time ot_{ws}, and its overall equipment effectiveness OEE_{ws}. Since these are all external parameters, the capacity can be pre-calculated. However, since they are key levers for adjustments in the resulting segmentation, it makes sense to model them explicitly.

3. Heuristic For Product Family-Oriented Product-Workstation Allocation Planning

Since heuristics are typically more applicable to complicated conditions than exact methods, it is often useful to create an accurate representation of the real problem and solve it heuristically, rather than looking for an exact solution to a more abstracted model [11]. Since integer programming problems are known to be hard to solve within reasonable time, the use of a heuristic also makes sense regarding the computing time for large real-world instances. Furthermore, heuristics with good comprehensibility increase the acceptance for their solution especially for practical problems like the one introduced in this paper. In the following, a greedy pareto heuristic procedure for the product family-oriented product-workstation allocation planning problem with restricted capacities is introduced as well as an outline of the required input data.

3.1 Procedure

The procedure for the proposed greedy heuristic is described by the following pseudo-code:

```
1   DO UNTIL all products are assigned:
2       CALCULATE $\emptyset\ Prio_{f,ws}$ per unassigned product family and available workstation
3       SELECT product family-workstation combination with highest $\emptyset\ Prio_{f,ws}$
4       FOR each prioritization level within selected combination in descending order:
5           FOR each product in descending order of required load $ct_{p,ws} \cdot d_p$:
6               IF available capacity $\geq$ required load of product $p$:
7                   Assign product to workstation: $x_{p,ws} = 1$
8                   Reduce available capacity by allocated workload
9       Block product family-workstation combination for further iterations
10      IF workstation has not enough available capacity:
11          Remove $ws$ from available workstations
12  END
```

The outlined procedure is iterated until all products are assigned to a workstation (line 1). First, the mean prioritization values $\emptyset\ Prio_{f,ws}$ are calculated for each product family-workstation combination with remaining unassigned products and available capacity (line 2). In line 3, the combination with the highest mean value is selected since it promises to involve many products with high prioritization values for the considered workstation. In the next step, the products of the selected product family are iteratively assigned to the selected workstation. Therefore, they are sorted in descending order of their prioritization value (line 4) and within each prioritization level, they are sorted in descending order of their required workload (line 5). Starting with the highest prioritization level, it is ensured that as many highly prioritized products as possible can be allocated to the selected workstation before its capacity may be exhausted. For each product it is checked if the available capacity of the considered workstation is still higher than its required load (line

6). If this is the case, the product is assigned to the workstation (line 7) and the remaining available capacity is reduced (line 8). If, on the other hand, the available capacity is not sufficient, the product remains unassigned and will be considered in further iterations. As soon as all products of the selected product family have been considered for the workstation assignment, the respective product family-workstation combination is blocked for all further iterations (line 9), since the workstation does not provide sufficient capacity for the remaining products of that family which are still to be assigned. Eventually, it is checked whether the selected workstation still has sufficient available capacity for further assignments in general (line 10). If this is not the case, it is removed from the list of available workstations. This procedure is repeated until all products have been assigned to a workstation.

Although the solution space may become very large for bigger problems, this heuristic is a simple and intuitive procedure to reach a good segmentation under consideration of allocation preferences and capacity restrictions. Its application is subject to some limitations regarding the capacity utilization. Firstly, it assumes a situation of scarce capacity so that in a case with large overcapacity it can result in a solution, where one workstation is assigned with two product families while another workstation is empty. In this case, the heuristic should be rerun with adjusted capacities. In the case of insufficient capacity, no feasible result is achieved since some products would remain unassigned. A splitting of the workload of one product does not take place. However, if splitting is necessary, it can be modelled by duplicating the corresponding product.

3.2 Required Input Data

For the application of the proposed heuristic, the following input data is required:

- Secondary demand for each product at each process d_p
- Cycle times per product and workstation $ct_{p,ws}$
- Product-process matrix with y_p
- Product-workstation-prioritization matrix with $Prio_{p,ws}$
- Product family assignment for each product $z_{p,f}$
- Workstation master data to calculate C_{ws}

The table with *secondary demand* d_p can be derived via the bill of material from the forecasted or historic overall demand of the finished products. *Cycle times* $ct_{p,ws}$ originate at best from time recordings or feedback data. However, planned times from the enterprise resource planning systems (ERP) can also be used with the knowledge in mind that they might be subject to inaccuracy. Multiplying d_p and $ct_{p,ws}$ gives the expected workload for each product at each workstation. A *product-process matrix* indicates which processes need to be passed by which product in the form of a binary matrix that holds the values for y_p. It is a standard tool for product family identification within the value stream method and if the information is not already available from the ERP, the matrix is either manually created or derived from transaction data [7]. The allocation preferences $Prio_{p,ws}$ are collected in the form of a *product-workstation prioritization matrix*. These are either assessed manually by experts or they can also be derived from historic production quantities mined from transaction data [7]. All products must be assigned to a *product family* which corresponds to the binary assignment parameter $z_{p,f}$. Eventually, the *workstation master data* must be given, including all information to calculate the available capacity C_{ws} of each considered workstation. These are the number of shifts s_{ws}, the operating time ot_{ws}, and the overall equipment effectiveness OEE_{ws}.

4. Evaluation

The proposed heuristic explores a new field in allocation planning. Thus, test data and assessment measures have not yet been established for a sound evaluation of the heuristic. Therefore, an extensive test set is generated, assessment measures are introduced, and comparison heuristics are used to rank the solution quality regarding the underlying solution space. All evaluation calculations are implemented in Python 3.7.6 in Visual Studio Code and run on a MacBook Pro with 1.4 GHz Quad-Core Intel Core i5 and 8 GB RAM.

4.1 Generation of test set

The test set is inspired by a dataset from an industry project concerning the value stream design for a battery cell production, on which the heuristic has also been applied. The test set includes 7776 problem instances that are generated with the use of random numbers and a full-factorial design combining the parameters listed in Table 2. Since the workstation allocation problem for one process is independent from every other process in a value stream, each test instance only consists of *one process*. The *number of workstations* is varied within 3, 6, 9, and 12, and for each workstation in the test instance, 10, 25, or 50 times as many products are generated. Assignment *prioritizations* are allocated randomly, such that each product has a first choice with a high prioritization value ($Prio_{p,ws}$= 10) and a second choice with a lower prioritization value ($Prio_{p,ws}$= 2). For all other workstations the prioritization is 0. The *demand* for each product is either constant 1000 pieces for each product or uniformly distributed between 800 and 1200 pieces. *Cycle times* are randomly determined within the intervals [10, 190], [50, 150], [90, 110]. The cycle times are distributed in three different ways. They can be equal for one product on every workstation, which represents the easiest case for the product-workstation allocation with respect to the capacity alignment. They can also be equal for one workstation with every product, which represents an environment of very different technology levels on parallel resources, such as an automated assembly station next to a manual one. In the third case, all cycle times are set independently. Based on the demand and the average cycle times of each product, an expected total load can be calculated. The *capacity* level of the workstations in the test instances is determined as the total load divided by an average utilization factor. This way, the available workstation capacity $\emptyset C$ is set to a value which allows to expect an average *utilization* of 80%, 90%, 95%, or 98% through the allocated products. The actual utilization in the eventual solution can, however, differ due to the varying cycle times. Products are grouped into *product families* either randomly or with the use of cluster algorithms based on their workstation prioritizations, which imitates the process similarity that products within a family typically have. For this purpose, the ward algorithm with Euclidean distances and the average-linkage approach with cosine distances is used. The number of generated product families is set to 2, 4, or 6.

Table 2: Parameters for test instance generation

Symbol	Description	Values			
$\|WS\|$	Number of workstations	3	6	9	12
$\|P\|$	Number of products per workstation	10	25	50	
-	Demand distribution	uniform	constant		
-	Cycle times distribution	by product	by workstation	independent	
δ_{ct}	Spread of cycle times [time units]	[10, 190]	[50, 150]	[90, 110]	
η	Average utilization of total capacity	80%	90%	95%	98%
$\|F\|$	Number of product families	2	4	6	
-	Cluster method for product family formation	ward	average-linkage	random	

4.2 Assessment of solution quality

The solution quality of the proposed pareto heuristic is assessed for the two objectives *total number of family-workstation assignments* FWA_{total} and *total prioritization* $Prio_{total}$ reached by the product-workstation assignments. To make the solutions of each problem instance comparable, FWA_{total} is set in relation with a lower bound LB and $Prio_{total}$ is set in relation with an upper bound UB. These bounds estimate the theoretically best possible values for the considered objective. As expressed in equation 12, the lower bound for FWA_{total} is either the number of product families or the number of workstations needed to cover the entire workload. For the workload estimation, the average cycle time of each product $\emptyset ct_p$ is multiplied with its demand d_p and summed up for all products. The division of this workload by the specified workstation capacity $\emptyset C$ of the test instance results in the minimum number of required workstations. The upper bound for $Prio_{total}$ is the sum of the highest prioritization value of each product as described by equation 13.

$$LB = \max\{\,|F|, \left\lceil \frac{\sum_{p \in P} \emptyset ct_p \cdot d_p}{\emptyset C} \right\rceil\} \tag{12}$$

$$UB = \sum_{p \in P} Prio_{p,max} \tag{13}$$

Based on these bounds, the evaluation results are given as ratios:

$$FWA\,[\%] = \frac{FWA_{total}}{LB} \cdot 100\% \tag{14}$$

$$Prio\,[\%] = \frac{Prio_{total}}{UB} \cdot 100\% \tag{15}$$

Therefore, values > 100% result for FWA [%] and values between 0% and 100% for $Prio$ [%]. The bound based ratios make different problem instances comparable and give an estimation on the quality of a solution. However, it is depending on the problem's structure how close to 100% a feasible solution can get for the two objectives. To get a better idea about this, three greedy comparison heuristics are implemented:

- *Random*: All products are sorted in descending order by average load and randomly assigned to a workstation with free capacity.
- *Max Prioritization heuristic*: All products are sorted in descending order by average load and iteratively assigned to the workstation with the highest prioritization, provided its capacity still allows it.
- *Min FWA heuristic*: All products are separated by product family and sorted in descending order by average load; starting with the family that has the highest total load, products are assigned block wise to the next free workstation.

With these simple heuristics, feasible solutions can be created for each instance, indicating how good one objective value can get when ignoring the second objective. Furthermore, the randomly created solution indicates how far the product-workstation allocation of the pareto heuristic is from a bad solution. With these comparative values, a classification in the solution space can take place, enabling a well-founded evaluation of the results.

4.3 Results

All solutions were calculated by the pareto heuristic within less than 10 seconds. On average, its runtime was 2.3 seconds. 6847 instances can be assessed, 929 instances lead to an infeasible solution in at least one of the applied approaches, which is due to problem instances that are too restrictive in terms of capacity. 695 of these unsolved instances are generated with the highest utilization expectation of η = 0.98.

The total results for the tested heuristics on all evaluable instances are illustrated in Figure 2. Averaged over all instances, the *pareto heuristic* reaches a prioritization ratio of $Prio$ [%] = 62, which is 13 percentage points behind the result of the *max prioritization heuristic*. A "natural low end" seems to be 44 percentage points lower at a prioritization ratio of $Prio$ [%] = 17. An even smaller gap can be observed for the FWA ratio. Here, the *pareto heuristic* reaches FWA [%] = 119 on average, while the *min FWA heuristic* reaches FWA [%] = 114. Compared to the average objective value of FWA [%] = 348 of the randomly generated solutions, the *pareto heuristic* is particularly strong in the target of reducing shared resources for clearly separatable value stream segments.

The achievable FWA [%] and thus the achievable number of shared resources depends above all on the scarcity of capacity (see Table 3), since this makes the problem more restrictive.

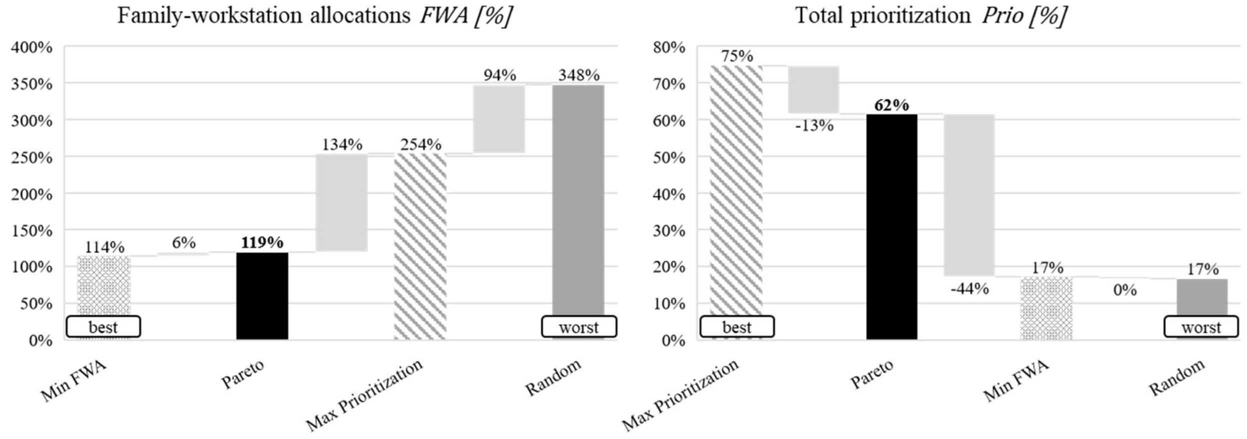


Figure 2: Average results for all heuristics in both objectives

If only those 4603 feasibly solved instances are considered for which the product families are formed using a clustering approach, the heuristic performs even better. In this case, the resulting values of the *pareto heuristic* are for both objectives almost as good as the comparison heuristic that ignores the other objective. The average FWA [%] is 117 compared to 113 of the *min FWA heuristic* and the average $Prio$ [%] is 73 compared to 75 of the *max prioritization heuristic*. Figure 3 illustrates the objective value pairs for these instances as well as their averages.

Table 3: Average FWA [%] for the pareto heuristic on all instances depending on the utilization factor η

η	FWA [%]
0.8	107.4
0.9	119.9
0.95	125.9
0.98	128.2

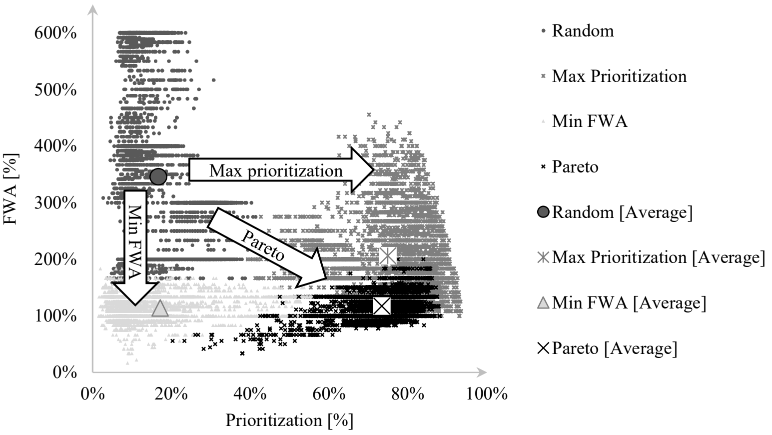


Figure 3: Results for all instances with clustering-based product families

5. Conclusion

This paper has introduced a pareto heuristic to support a product family-oriented segmentation under consideration of the available capacities as well as workstation preferences. The mathematical formulation of the multi-objective optimization problem with its restrictions is outlined, the heuristic procedure is presented, and the required business data is described. For the evaluation, an extensive test set is generated, assessment measures are introduced, and comparison heuristics are implemented. The evaluation shows, that the pareto heuristic performs strong for both objectives, especially for the target of reducing shared resources in a value stream design. As expected, even better results are achieved, if the product family affiliation correlates with the workstation preferences of the products. The value stream method provides to identify product families based on similarity in their required process steps and equipment. Therefore, high correlations between product family affiliation and workstation preferences should be observable in a well-executed value stream design. The limitations of the heuristic are in its applicability for scenarios of insufficient capacity, large overcapacity, or if the high workload of one product would be required to be split. However, project-specific conditions can easily be modelled either via an adaption of the heuristic or via the given input data. The resulting product-workstation assignments of the pareto heuristic provides practitioners with a good and well-founded segmentation as a basis for the next steps in the value stream design that builds on it.

Acknowledgements

The research leading to these results has received funding from the German Federal Ministry of Education and Research under grant agreement number 02L19C150 (KompAKI).

References

[1] Serrano, I., Ochoa, C., Castro, R.D., 2008. Evaluation of value stream mapping in manufacturing system redesign. International Journal of Production Research 46 (16), 4409–4430.

[2] Lugert, A., Batz, A., Winkler, H., 2018. Empirical assessment of the future adequacy of value stream mapping in manufacturing industries. Journal of Manufacturing Technology Management 29 (5), 886–906.

[3] Braglia, M., Carmignani, G., Zammori, F., 2006. A new value stream mapping approach for complex production systems. International Journal of Production Research 44 (18-19), 3929–3952.

[4] Urnauer, C., Gräff, V., Tauchert, C., Metternich, J., 2020. Data-Assisted Value Stream Method. WGP Jahreskongress.

[5] Rother, M., Shook, J., 1999. Learning to see: Mapping the value stream to add value and eliminate waste. Lean Institute, Cambridge.

[6] Erlach, K., 2013. Value Stream Design. Springer, Berlin.

[7] Urnauer, C., Metternich, J., 2022. Product family identification based on data analytics. Procedia CIRP [in press].

[8] Duggan, K.J., 2013. Creating mixed model value streams: Practical lean techniques for building to demand, 2nd ed. CRC Press Taylor & Francis Group, Boca Raton, London, New York.

[9] Gaida, M., Günther, U., Wilsky, P., Riedel, R., 2020. Bildung von Produktfamilien als Planungsgrundlage auf Basis von Clusteralgorithmen. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (3), 111–114.

[10] Erlach, K., Böhm, M., Gessert, S., Hartleif, S., Teriete, T., Ungern-Sterrnberg, R., 2021. Die zwei Wege der Wertstrommethode zur Digitalisierung: Datenwertstrom und WertstromDigital als Stoßrichtungen der Forschung für die digitalisierte Produktion. Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (12), 940–944.

[11] Reeves, C.M. (Ed.), 1993. Modern heuristic techniques for combinatorial problems, 1. publ ed. Halsted Press, New York.

Biography

Christian Urnauer, M. Sc. (*1990) is Research Associate at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt. He holds a Master of Science in Industrial Engineering which he obtained at the Technical University of Darmstadt. His research interests are in value stream mapping and data analytics.

Anna-Maria Larem, M. Sc. (*1994) is Manufacturing Trainee in the Graduate Specialist Program at Bosch Rexroth AG. She holds a Master of Science in Industrial Engineering which she obtained at the Technical University of Darmstadt. Her research interests are in lean production and data analytics.

Prof. Dr.-Ing. Joachim Metternich (*1968) is head of the Production Management division at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt. He is spokesman of the medium-sized businesses digital competence centre Darmstadt; he has been president of the International Association of Learning Factories (IALF) from 2015 to 2021 and he is a member of the scientific society for production technology (WGP) as well as scientific partner of the Conference on Production Systems and Logistics (CPSL).

3rd Conference on Production Systems and Logistics

Cross-Company Routing Planning: Determining Value Chains In A Dynamic Production Network Through A Decentralized Approach

Julia Markert[1], Dominik Saubke[1], Pascal Krenz[1], Lothar Hotz[2]
[1]*Helmut-Schmidt-University, Hamburg, Germany*
[2]*Hamburger Informatik Technologie-Center, Hamburg, Germany*

Abstract

Demand-based, local production will gain relevance in the context of sustainability and circular economy. One way to implement local value creation is through establishing highly dynamic networks that consolidate the competencies of regional manufacturers. Consequently, the structure of the value chains needs to be determined ad hoc dependent on demand. This is a rather challenging task due to the dynamics within such networks and the flat hierarchies. Traditionally, value chains are defined and controlled in a centralized form by a lead firm or a separate stakeholder (e.g. Intermediary, Broker). However, to accommodate the dynamics of demand and the increasing complexity of products, we propose a decentralized form of coordination. The basic idea is to upscale Routing Planning, used in Process Planning, to a network level. Meaning instead of a centralized instance within a company defining the production steps, the stakeholders will collaboratively determine the cross-company Routing Plan, effectively building the value chain. Thus, the accumulated experience and knowledge of all stakeholders can be utilized to efficiently fulfil current customer demand, since the value chain will be executed by the same stakeholders that created it. But in order to coordinate the sequencing of operations by multiple stakeholders, suitable methods need to be implemented. We look at a strategy to facilitate such a collaboration between companies and demonstrate one possible technical implementation based on AI planning using Planning Domain Definition Language (PDDL).

Keywords

Production network; Process Planning; Collaborative planning; Value chain; AI planning

1. Introduction and Motivation

While global value creation with division of labour is prevalent nowadays, in recent years the concept of producing locally in small networked sites has gained an increasing amount of attention [1–3]. There is more and more research dealing with new, demand-based local production principles, such as Urban and (Re-) Distributed Manufacturing. These concepts aim to reach the goals of a sustainable production and fulfil customers' growing desire for increasingly individualized products [4,5,1]. Larsson even went so far as to state that "there will be a need to re-shape entire value chains and a large share of the corporate landscape" [1]. Additionally, the Covid-19 pandemic has accelerated research interest in local production as an exploding need for medical equipment has shown the risky dependence on global value chains [6,2]. The local production of masks was able to alleviate the shortage, effectively demonstrating an important advantage of local production - resilience [2].

DOI: https://doi.org/10.15488/12135

One challenge of local production is being able to adapt to demands changing fast due to increasing product complexity and individualization. To achieve a high product variety and small batch sizes, a multitude of production processes has to be covered by local stakeholders. Multiple companies have to collaborate in networks in order to combine their competences and to fulfil these requirements [7]. However, constructing highly flexible, adaptable value chains in local networks calls for appropriate strategies, because unlike in mass production, long-term collaborations with carefully selected partners and high degrees of insight into suppliers' production are not feasible. Instead a high level of dynamics develops within the structures of local value creation, which in turn reduces transparency. For that reason, we will discuss a decentralized strategy for creating value chains in the described environment and one possible technical solution based on AI planning (Artificial Intelligence) as it "has been successfully applied for decades in several areas" [8].

2. Theoretical Background

This paper will talk about the idea to use principles from Operations Planning and Scheduling (OPS), more specifically Process Planning (PP), in order to create value chains in a decentralized network of local small and medium-sized enterprises (SME). Therefore, Chapter 2.1 will discuss existing concepts for the planning of value chains in networks while Chapter 2.2 will give some fundamentals regarding Process Planning.

2.1 Network Planning

The network's status as a very important and most modern organisation form of producing companies [9] is in line with the increasing trend of outsourcing, which basically means transferring tasks and processes to an outside party [10]. Over the last decades, outsourcing has become a standard practice in many companies for various reasons, such as decreasing costs and lowering the vertical range of manufacture [9,10]. Industrial production today is shaped by a high degree of often international division of labour. That means, companies producing physical goods build a network of specialized suppliers and other firms to reach their production goals. The interaction and communication along the value chain is facilitated by, e.g., product or industry standards, modularization and platform technologies. However, when parts or even just production steps are outsourced, costs and effort for coordinating production will increase [9]. Usually the focal company selling the end product will procure the needed parts itself by distributing the necessary orders to suppliers. Having a network of suppliers, with whom a company works regularly, as well as the mentioned interaction mechanisms can therefore make this process more efficient. This hierarchical concept works well when the production program is planned months or years ahead, but comes to its limits when the production needs to be dynamic with individualized, everchanging products in small batch sizes or single product orders [11,12].

This paper targets a decentralized, local production as another form of value creation next to global production with division of labour. In such a production aiming to serve local demands ad hoc, a new dynamic develops. Because the sales market in a local production is small, the batch sizes are as well, but the product variety is very high. This leads to fluctuating demands in a short time while production has to be handled by comparatively few manufacturers, since resources are limited by the regionality of local production. In order to adequately handle this dynamic in combination with the mentioned parameters (local, on demand, high product variety, fewer producers) new solutions are needed.

One of these is a concept introduced in the late 90s by Schuh et al. [7] - the Virtual Factory[1]. Several companies come together to form a larger, Virtual Factory to easier handle short term production in small batches. There is a stable and a dynamic component involved in such a Virtual Factory [7]. The relationships

---

[1] This is not to be confused with the virtual factory concept as introduced in 1993 by Onosato and Iwata [13] that focuses on modelling manufacturing. More recent information on this strain of virtual factory research can be found, e.g., by Debevec et al., who dealt with simulating production processes in SMEs [14], or Yildiz et al., who wrote about a digital twin-based virtual factory [15].

within the network should be stable, as they are in a typical production network as well, while the processing of orders is done dynamically through "activating" networks for each order [7]. Schuh et al. [7] defined six intercorporate services to manage the Virtual Factory network, one of them being the "broker". The "broker" is usually an independent party responsible for – among other tasks such as acquiring customers – distributing the competences of the companies to where they are needed while roughly defining prices [7]. The broker may also initiate the so called "active networks", meaning essentially the value chains that will go into action for a specific order or project [7].

2.2 Operations Planning and Scheduling and Process Planning

Operations Planning and Scheduling (OPS) essentially describes all the organisational functions that need to be fulfilled to bring a product design into production and to the customer [16,17]. Process Planning (PP) is one part of OPS, the other part is Production Control (also referred to as Production Planning and Control or PPC). PP encompasses the planning necessary to ensure the economically sound manufacturing in line with production requirements and it is typically done once without reference to a specific order [18,19]. As shown in Figure 1, the tasks of PP can be divided into three categories, short-, medium- and long-term as introduced by Eversheim [19] and referenced by, e.g., Bauernhansl and Spur [18,20]. PPC on the other hand includes the measures needed to fulfil actual orders based on the results of PP [17].

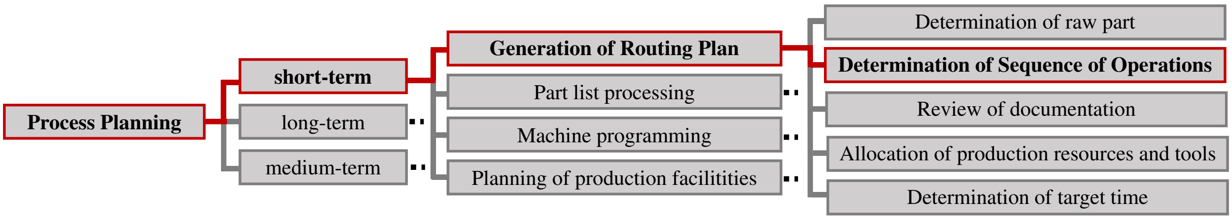


Figure 1: The structure of Process Planning (with selected details) [18,19]

One result of PP is the Routing Plan (RP), which is described as "a plan defining the sequence of operations to be performed in order to produce a part or an assembly, i.e. the product's path through the production process" [16]. Besides the production steps in the right order, a RP will include information regarding the type of raw part to start with, production resources and tools as well as the target time for each step [19]. For this paper, the focus will be RP, especially the sequence of operations, marked by red borders in Figure 1.

3. Concept: Implementing Operations Planning and Scheduling on a Network Level

Modern local production leads to a necessity to manufacture dynamically and on-demand to be able to fulfil customer needs in times of individualization without having large amounts of products in warehouses available to do so. The dynamic nature of networked production also requires a cross-company organisation that will be able to keep up the fast pace of small-batch and single product orders. Since building value chains spanning multiple companies bears many similarities to building an in-house RP (refer to Chapter 2.2), it can be useful to apply the methods and tasks of RP to a network/cross-company level. The extensive knowledge and research on OPS and specifically RP already available can then be used as a baseline for adapting existing concepts and deriving new ones for the given context.

3.1 Approach for the decentralized determination of value chains

In a hierarchical network (as introduced in Chapter 2.1) the lead firm typically uses its network of suppliers to coordinate value creation by acquiring materials and parts as well as outsourcing production. Two examples for this kind of centralized approach are Wintelism, which is described as the "extensive outsourcing of component production to enable industrial structures to become less vertically and more horizontally integrated" [21], and Sturgeon's modular production networks [22]. In addition, the planning departments of each company involved in the network generate in-house RPs, meaning they determine the sequence of operations within their own company and allocate the production resources (Figure 1) [19].

In centralized networks it is the lead firm's task to configurate the cross-company value chains. There are usually relatively stable planning horizons that facilitate mid- to long-term planning. However, when dealing with dynamic, heterarchical and decentralized networks there are no steady planning periods. As a result, the generation of value chains needs to happen faster and adapt dynamically, so that even single product orders can lead to entirely new value chains. In order to move away from the more permanent structures of value chains, the network can be viewed as one virtual company, similar to Schuh's Virtual Factory [7] (refer to Chapter 2.1), with the SMEs representing production resources. This also serves the depiction of the heterarchical structure of a decentralized network. Using this analogy, value chains correspond to the RPs introduced in Chapter 2.2 with the difference that they now function as cross-company RPs (ccRP). This view and the introduction of ccRPs should allow for better reactivity in local, decentralized networks.

However, other than in the creation of traditional RPs, there is no superior central entity that knows all the details of the production of all the companies involved in the network. That is because there is often a lot of experience involved in writing RPs. This is categorized as tacit knowledge, i.e., it is not formalized and thus difficult to share [23]. Nonaka calls the process of making tacit knowledge sharable by turning it into explicit knowledge "externalization" [23]. Externalizing the knowledge needed for common routing planning from individuals of the companies in order to create such a central entity would be time-consuming and expensive, given that the involved companies would even agree to share their knowledge to the required extent. The aforementioned dynamics of processes and structures further add to the difficulties of externalization. As a result, when working with a heterarchical, dynamic network, it is highly unlikely that a central entity has the knowledge needed to determine the best possible ccRPs for all involved companies or even any at all.

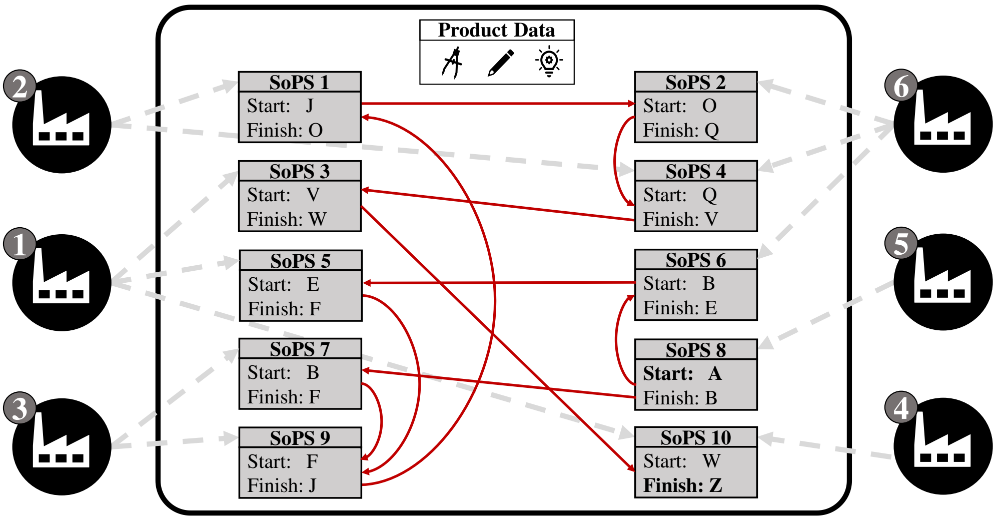
We therefore suggest to avoid the externalization process altogether by **taking a decentralized approach to ccRP**. That means instead of gathering knowledge from the companies, that knowledge stays where it is, but is still used to create the ccRPs. According to Krenz, the complexity of the knowledge that is to be shared needs to be reduced, so that the stakeholders of the network are able to collaborate [12]. In summary, the concept idea we propose in this paper seeks to create value chains in a decentralized way through leaning on the principles of OPS and PP and upscaling them to a network or cross-company level.

This is also an important point of differentiation from Schuh et al.'s concept of a Virtual Factory, as described in Chapter 2.1. While it also involves a decentralized, local network, Schuh et al. still assign various functions or centralized roles to distinct stakeholders of the network such as the "broker", while the ccRP concept relies on decentralized collaboration of the companies within the network instead.

3.2 Introduction of a possible strategy

In order to implement decentralized ccRP in dynamic networks, it is necessary to find a way for process planners from all the network companies to contribute to a cumulative plan in a simple and fast manner. The simplest form of collaboration would be **experts discussing and developing a plan together**. This however bears several issues such as the group size getting too large for useful discussions, experts being invited unnecessarily and the need for a skilled and neutral facilitator. Overall this approach would incur unreasonably high costs and consume a lot of time compared to the central approach while having presumably little advantages. In conclusion, the decentralized approach needs to allow planners to be able to work independently from one another, ideally in regards to time and location, also called **asynchronous communication** [24], and it needs to be well structured and efficient. This is typically prevalent in cross-company collaborations [25].

For ccRP the goal is to sequence the economically best value chains while using the planners' expertise without the need for direct communication between companies. For this strategy proposal we, on the one hand, drew inspiration from **collaboration in communities**, where everyone may contribute their expertise. But on the other hand, since ccRPs are a planning problem, we also relied on research on automated planning and scheduling (AI planning), namely **using world states for computing plans instead of defining processes** [26]. The focus is to develop an approach **in which process planners can work together without directly communicating**. Instead they contribute their knowledge and competence in such a form, that an intelligent and autonomous planning system (example in Chapter 3.3) can generate an optimized ccRP and thus create a cohesive value chain. That strategy is visualized in Figure 2. The premise is a dynamic, heterarchical network, where product development can be done by a company, an engineer without own production capacities or a community of designers, etc. There is no central entity distributing/outsourcing production steps. Instead, the Product Data is provided to a network of process planners that will then derive the production steps their own company can do and enter them into a network system. However, they do not actually describe what they are doing in production or how they are doing it. Instead the only thing entered into the system is how the product changes, i.e., the planners describe the product state they can start production with and the product state at the end of their company's part in the production process, e.g. bar stock - gear shaft (unhardened). They therefore contribute a **Sequence of Product States (SoPS)**.



SoPS = Sequence of Product States

Figure 2: Decentralized, cross-company Routing Planning and value chain creation

In Figure 2 those product state descriptions are represented by letters, i.e., the starting state of SoPS 1 is "J" and the finishing state is "O". These descriptions are used to find interfaces between the companies, e.g., between SoPS 1 ("O" as the finishing state) and SoPS 2 ("O" as the starting state). The possibility of branching production sequences, i.e., sequences that can be processed in parallel until the parts are joined together, e.g., in assembly, is not visualized in Figure 2 for reasons of clarity and comprehensibility. In such a case, the start or end state would be the combination of two states, e.g., "C and D". The necessary formalization of the state descriptions is discussed in Chapter 4. According to Dietl, companies will delimit their value creation activities in such a way that as little implicit knowledge as possible needs to be transferred in between steps [27]. In other words, the steps of value creation are separated in a way that the knowledge exchange at the interfaces is reduced to a minimum transfer of external knowledge [27]. The stakeholders integrate their implicit knowledge into the product and thus transfer it by changing the product state [27,12]. When all steps are entered, several viable paths are calculated (for reasons of clarity and comprehensibility only two are shown here) and one or more optimized paths (depending on the system) are determined. Those are the ccRPs available for PPS (refer to Chapter 2.2), i.e., the implementation of the value chains. When feeding the system with further information such as capacity, price, transport distances and number of involved companies, the most fitting ccRP at a given time can be chosen according to selected criteria. This is however a separate topic and not further discussed in this paper. Overall, the presented strategy would allow the ccRPs to develop in a self-organizing process in which the network companies choose the interfaces. Therefore, this strategy for configurating value chains can be considered heterarchical.

This may lead to considerable benefits in terms of network dynamics and reactivity. There are several viable options for a value chain available, so one alternative can be chosen to be used for actual production. Additionally, the chosen path can easily be changed mid-production in case of unforeseen events since alternatives have already been determined beforehand. When several companies are capable of performing the same steps, it creates redundancy and therefore resilience. This enables the network to adapt in case, e.g., capacity issues occur and one company is not able to deliver. The strategy also has the advantage that the process planners can work entirely independent of one another, they do not have to spend time to understand each other's work or to standardize plans cross-company. They can use their expertise without sharing it, which is especially important when sensitive, inside knowledge of companies is involved. Information necessary for planning is fed to the system to be processed instead of being shared in the network community.

3.3 Technological approach for the implementation of decentralized ccRP

3.3.1 AI Planning using the Planning Domain Definition Language (PDDL)

The presented strategy for decentralized planning illustrates that this problem can be solved with the methods and systematics of general decentralized planning principles. Decentralized networks of production units already require a high planning effort with a small number of stakeholders and less complex products, such systems exhibit multitude branching, which can only be evaluated with computer support [28].

The so-called research area of AI Planning within the technology field of AI deals with the solving of complex planning tasks with high numbers of possible solutions. Solving these problems involves the change of states through sets of *action*s. The initial situation is a problem with an initial state and a final state, the *goal*. Hereby details of the change or effect are not considered, i.e., actions are only described through input (a *precondition*) and output (an *effect*) states. In our use case, this leads to a special advantage, since company-internal knowledge of concrete production steps is not revealed in detail. [26,29]

The implementation technology pursued here is PDDL. With the introduction of PDDL in 1998, inspired by the STRIPS (Stanford Research Institute Problem Solver) formalism, a standardization of AI planning languages was developed [30,31]. PDDL is considered to currently be one of the most common solutions to implement AI planning [30] and thus suitable to address the underlying planning problem of ccRP. The solution to a planning problem, the fulfilment of target conditions, is considered to be a concatenation of actions, i.e., a concatenation of Sequences of Product States (SoPS) as shown in Figure 2 (Chapter 3.2). Those SoPS are entered into the system by the individual stakeholders in the production network. As already described, they can contain different numbers of production and process steps. Each SoPS contains a change of the state. Furthermore, following the principle of PDDL, each SoPS will be modelled as an action and will be assigned a precondition and an effect. By matching this information, the SoPS are aligned to find a path to the end-state (*goal*). In the next section a PDDL implementation approach is presented. [26,28,32]

3.3.2 Approach for the implementation of PDDL in the given context

As an example, in Figure 3, we provide a PDDL model for the SoPSs depicted in Figure 2 as a domain model and a problem model, respectively. The domain model describes the rules of RP as predicates and actions, while the problem model describes the actual task at hand, e.g., to compute the plan from subproduct "A" to subproduct "Z". Another task would be to compute the plan from subproduct "B" to subproduct "F". The intermediate production results of each SoPS are named SUBPRODUCTS and modelled as objects. One action (named SoPSx, where x is the numbering of the SoPS in Figure 2) models one SoPS with a certain SUBPRODUCT as precondition ("Start") or effect ("Finish"). Furthermore, a specific SUBPRODUCT is marked as initial state (i.e., "A") and another as goal (i.e., "Z").

```
(1) Problem Model in PDDL

Objects:
Subproducts: A, B, E, F, J, O, Q, V, W, Z
In PDDL:
(:objects A, B, E, F, J, O, Q, V, W, Z)

Initial State:
SUBPRODUCT(A) is true, i.e.,
A is the start of the process

In PDDL:
(:init (SUBPRODUCT A))

GOAL specification:
SUBPRODUCT(Z) is true, i.e., Z shall be produced

In PDDL:
(:goal (SUBPRODUCT Z))
```

```
(2) Domain Model in PDDL

Predicates:
SUBPRODUCT(x) - true iff x is a subproduct
In PDDL:
(:predicates (SUBPRODUCT ?X))
; whereby ?X represents a variable

Action:
Description: Each SoPS of Figure 3.2
is represented by one action.

Precondition is the „Start" subproduct
Effect is the „Finish" subproduct

In PDDL:
(:action SoPS1                         (:action SoPS6
   :precondition (SUBPRODUCT J)           :precondition (SUBPRODUCT B)
   :effect (SUBPRODUCT O))                :effect (SUBPRODUCT E))
(:action SoPS2                         (:action SoPS7
   :precondition (SUBPRODUCT O)           :precondition (SUBPRODUCT B)
   :effect (SUBPRODUCT Q))                :effect (SUBPRODUCT F))
(:action SoPS3                         (:action SoPS8
   :precondition (SUBPRODUCT V)           :precondition (SUBPRODUCT A)
   :effect (SUBPRODUCT W))                :effect (SUBPRODUCT B))
(:action SoPS4                         (:action SoPS9
   :precondition (SUBPRODUCT Q)           :precondition (SUBPRODUCT F)
   :effect (SUBPRODUCT V))                :effect (SUBPRODUCT J))
(:action SoPS5                         (:action SoPS10
   :precondition (SUBPRODUCT E)           :precondition (SUBPRODUCT W)
   :effect (SUBPRODUCT F))                :effect (SUBPRODUCT Z))
```

Figure 3: (1) Problem description that has a specific start and goal; (2) Domain model describing the SoPS as actions

Given this domain and problem model to PDDL, it will compute a plan (or multiple plans) that takes the action definition into account, i.e., it would compute a value chain. According to the dependencies, two paths are possible in our example, one from SoPS8 via SoPS7 to SoPS9 and one via SoPS6 and SoPS5 (Figure 4). In case of branching production sequences (as mentioned in Chapter 3.2) the precondition of an action would consist of two or more subproducts, e.g., `(:action SoPSWithAssembly :precondition (and (SUBPRODUCT W1) (SUBPRODUCT W2)) :effect (SUBPRODUCT AssembledZ))`.

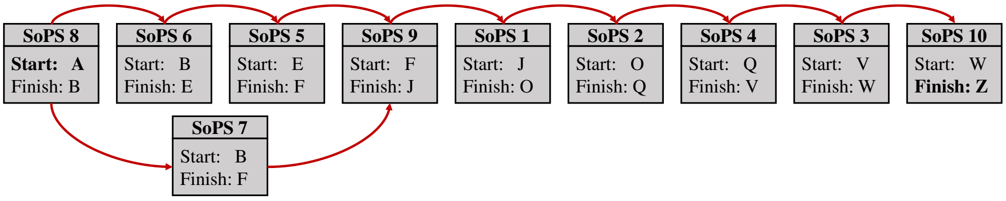


Figure 4: Possible paths in the given example according to their dependencies

3.3.3 Assembling ccRPs using Heuristic Methods

The core of PDDL is to use a heuristic search algorithm [33] for computing appropriate plans as described, i.e., PDDL implements such an algorithm, e.g., A*. Alternatives to PDDL, which use heuristic search, are distributed AI algorithms based on agents (agent-based algorithms), where each SoPS would be considered as one agent and by following a negotiation process the agents would compute a plan. Other alternatives would be evolutionary algorithms (EA) which would compute plans through creating plans by mutating them randomly and evaluating those through domain related criteria such as order of SoPS (selection).

4. Critical Analysis of the Proposed Concept

When analysing the presented concept concerning its feasibility in a real application, the criteria must be of **strategic**, **operational** or **technical** nature. In terms of strategic application, the concept of a decentralized collaborative OPS seems particularly suited to the requirements of new production structures, with increasingly blurred physical boundaries between companies.

In operational terms, it is to be mentioned that the strategy presented in Chapter 3.2 is still in the early stages and only an outline. There are several issues to be detailed and addressed in future research to develop the strategy into an implementable concept, since at this point it only describes a way to get to the description of technically feasible value chains. Before one of those can be chosen for implementation, they need to be evaluated using criteria like timing and cost. It is also thinkable to include other variables for choosing one of the plans, such as energy efficiency of production, transport distances, etc. Additionally, the "perfect" value chain may still not be implementable at any given time due to capacity and scheduling issues. RP is only one step in the interface between design and production, our strategy needs to be embedded in a holistic concept covering all of OPS in order to work. Other issues to be considered in future research are, e.g., data storage and security, but also issues of motivation, incentives to participate, building trust in a network, etc.

When critically examining the technical implementation, the main focus is on the description and formalization of the SoPS, especially since the system needs to recognize when two planners are describing the same state in different words. Precondition, parameter and effect must use a uniform language in order to be linked with each other. An approach to formalization could be the use of Natural Language Processing (NLP). This is a solution for a further improved human-computer interaction and a common tool in the field of AI. In addition, an investigation and more exact statement in regards to the critical size of the network seems meaningful, especially for concrete decisions in technology selection (PDDL Toolboxes, Heuristic Search Algorithm, etc.). In general, the topic of quality management in such a system is likely to face new challenges, which are already evident here at the beginning when evaluating planning solutions. Also, when selecting and applying different methods from the field of heuristic search, attention must be paid to their result orientation. Depending on the modelled optimization criteria, the heuristic search will provide optimized solutions. In our example we focused on dependencies between the SoPS. Other criteria could include the ones mentioned above on the operational level. Thus, when considering value chains, the chosen termination criterion will have a direct impact on quality, cost, sustainability and customer satisfaction.

5. Outlook

In conclusion, using decentralized, cross-company Routing Planning in order to dynamically create value chains for local networks while relying on existing knowledge within the involved companies seems to be a promising approach. For one it draws on decades of research on Operations Planning and Scheduling as well as the often extensive expertise of experienced process planners, but it also combines these with state-of-the-art information technologies from the fields of AI planning and Heuristics. In our approach the knowledge stays with the experts, but is still utilised and arranged by an intelligent system in order to generate flexible value chains in a decentralized way.

This opens up several possibilities for future research. Firstly, on an operational level, as mentioned in the critical analysis, the strategy of using changes in the product state described by a network of process planners to find a Routing Plan should be developed further and with more detail. Following that, the second part of Operations Planning and Scheduling, Production Control (also known as Production Planning and Control), could be addressed. In this paper we looked at using Process Planning on a cross-company level and decentralizing it. The next logical step would be to explore whether or not Production Control principles are also applicable and beneficial to be used in networked production.

Furthermore, the technological approach discussed in Chapter 3.3 will be optimized through an ongoing project to implement a decentralized, local production network in the Greater Hamburg area. Going into more detail and possibly simulating a test run will be a next step. It would also be expedient to address the issue of comparing and assessing the cross-company Routing Plans created by the heuristic based on selected criteria. Additionally, using PDDL and heuristic methods is just one possible way to go. It might be interesting to explore other technologies that could be suitable for the presented problem.

Acknowledgements

This research is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [project: Production Next Door].

References

[1] Larsson, M., 2018. *Circular business models: Developing a sustainable future*. Palgrave Macmillan, Cham.

[2] Peters, A., Guitart, C., Pittet, D., 2021. Addressing the global challenge of access to supplies during COVID-19, in: Hadi Dehghani, M. (Ed.), *Environmental and Health Management of Novel Coronavirus Disease (COVID-19).* Elsevier Science & Technology, San Diego, pp. 419–441.

[3] Spath, Dieter (Hrsg.), Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., Schlund, S., 2013. *Produktionsarbeit der Zukunft - Industrie 4.0: Studie*. Fraunhofer Verlag, Stuttgart, 150 pp.

[4] Kohtala, C., 2015. Addressing sustainability in research on distributed production: an integrated literature review. *Journal of Cleaner Production 106*, 654–668.

[5] Krenz, P., Stoltenberg, L., Markert, J., Saubke, D., Redlich, T., 2022. The Phenomenon of Local Manufacturing: An Attempt at a Differentiation of Distributed, Re-distributed and Urban Manufacturing, in: Andersen, A.-L. et al. (Eds.), Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems. *Proceedings of the 8th Changeable, Agile, Reconfigurable and Virtual Production Conference (CARV2021) and the 10th World Mass Customization & Personalization Conference (MCPC2021)*, 1st ed. 2022 ed. Springer International Publishing; Imprint Springer, Cham, pp. 1014–1022.

[6] Hartig, S., Duda, S., Hildebrandt, L., 2020. Urgent need hybrid production - what COVID-19 can teach us about dislocated production through 3d-printing and the maker scene. *3D printing in medicine* 6 (1), 37.

[7] Schuh, G., Millarg, K., Göransson, Å., 1998. *Virtuelle Fabrik: Neue Marktchancen durch dynamische Netzwerke*. Hanser, München, 186 pp.

[8] Vallati, M., Kitchin, D., 2020. *Knowledge Engineering Tools and Techniques for AI Planning*, 1st ed. 2020 ed. Springer International Publishing; Imprint Springer, Cham, 277 pp.

[9] Schuh, G., 2006. *Produktionsplanung und -steuerung: Grundlagen, Gestaltung und Konzepte*, 3., völlig neu bearb. Aufl. ed. Springer, Berlin, Heidelberg, 876 pp.

[10] van Weele, A.J., Eßig, M., 2017. *Strategische Beschaffung: Grundlagen, Planung und Umsetzung eines integrierten Supply Management*. Springer Gabler, Wiesbaden, 631 pp.

[11] Buxbaum-Conradi, S., 2018. *Global and local knowledge dynamics in an industry during modular transition: A case study of the Airbus production network and the Aerospace Cluster in HH*, Northern Germany. Diss., HH.

[12] Krenz, P., 2020. *Formen der Wissensarbeit in einer vernetzten Wertschöpfung*. Diss., Hamburg.

[13] Onosato, M., Iwata, K., 1993. Development of a Virtual Manufacturing System by Integrating Product Models and Factory Models. *CIRP Annals 42 (1)*, 475–478.

[14] Debevec, M., Simic, M., Jovanovic, V., Herakovic, N., 2020. Virtual factory as a useful tool for improving production processes. *Journal of Manufacturing Systems 57*, 379–389.

[15] Yildiz, E., Møller, C., Bilberg, A., 2020. Virtual Factory: Digital Twin Based Integrated Factory Simulations. *Procedia CIRP 93*, 216–221.

[16] CIRP - The International Institution for Production, 2020. *Dictionary of Production Engineering III – Manufacturing Systems Wörterbuch der Fertigungstechnik III – Produktionssysteme Dizionario di Ingegneria della Produzione III – Sistemi di produzione*. Springer Vieweg, Berlin, Heidelberg.

[17] Minolla, W., 1975. *Rationalisieren in der Arbeitsplanung: Schwerpunkt Organisation*. Diss.

[18] Bauernhansl, T. (Ed.), 2020. *Fabrikbetriebslehre 1: Management in der Produktion*, 1. Aufl. 2020 ed. Springer Vieweg, Berlin, Heidelberg, 388 pp.

[19] Eversheim, W., 1989. *Organisation in der Produktionstechnik: Arbeitsvorbereitung*, Zweite, neubearbeitete Auflage ed. Springer Berlin Heidelberg; Imprint; Springer, Berlin, Heidelberg.

[20] Spur, G., Stöferle, T., 1994. *Fabrikbetrieb*. Hanser, München, 385 pp.

[21] Hart, J.A., Kim, S., 2002. Explaining the Resurgence of U.S. Competitiveness: The Rise of Wintelism. *The Information Society 18 (1)*, 1–12.

[22] Sturgeon, T.J., 2002. Modular production networks: a new American model of industrial organization. *Industrial and Corporate Change 11 (3)*, 451–496.

[23] Nonaka, I., 1994. A Dynamic Theory of Organizational Knowledge Creation. *Organization Science 5 (1)*, 14–37.

[24] Li, W.D., Qiu, Z.M., 2006. State-of-the-art technologies and methodologies for collaborative product development systems. *International Journal of Production Research 44 (13)*, 2525–2559.

[25] Jiang, P., Shao, X., Qiu, H., Li, P., 2008. Interoperability of Cross-organizational Workflows based on Process-view for Collaborative Product Development. *Concurrent Engineering 16 (1)*, 73–87.

[26] LaValle, S.M., 2006. Planning Algorithms. Cambridge University Press, Cambridge.

[27] Dietl, H., 1993. *Institutionen und Zeit*. Zugl.: München, Univ., Diss., 1991. Mohr, Tübingen, 246 pp.

[28] Nägele, L., Schierl, A., Hoffmann, A., Reif, W., 2018. Automatic Planning of Manufacturing Processes using Spatial Construction Plan Analysis and Extensible Heuristic Search. *Proceedings of the 15th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2018) - Volume 2*, pages 576-583.

[29] Russell, S.J., Norvig, P., 2016. *Artificial intelligence: A modern approach*, Third edition, Global edition ed. Pearson, Boston, Columbus, Indianapolis, 1132 pp.

[30] Haslum, P., Lipovetzky, N., Magazzeni, D., Muise, C., 2019. *An introduction to the Planning Domain Definition Language*. Morgan & Claypool, San Rafael, California, 169 pp.

[31] Wodecki, A., 2019. *Artificial Intelligence in Value Creation*. Springer International Publishing, Cham.

[32] Tencent Research Institute, Internet Law Research Center (CAICT), Tencent AI Lab, Tencent open platform, 2021. *Artificial Intelligence*. Springer Singapore, Singapore.

[33] Edelkamp, S., 2012. *Heuristic search: Theory and applications*. Morgan Kaufmann, Waltham, MA, 836 pp.

Biography

Julia Markert, M.Sc. (*1996) has been a research assistant at the Institute of Production Engineering (LaFT) within the Department of Mechanical Engineering of Helmut-Schmidt-University in Hamburg since 2021. She has a background in Production Engineering and her current research focus is on the creation and coordination of dynamic value chains in local production networks.

Dominik Saubke, M.Sc. (*1989) has been a research assistant at the Institute of Production Engineering (LaFT) within the Department of Mechanical Engineering of Helmut-Schmidt-University in Hamburg since 2021. With several years of practice experience in Product Engineering and Product Development his current research focus is production-readiness in collaborative crowd design processes.

Pascal Krenz, Dr.-Ing. (*1983) has been a research assistant at the Institute of Production Engineering (LaFT) within the Department of Mechanical Engineering of Helmut-Schmidt-University in Hamburg since 2011. He leads the research area on local production at the Institute and related research projects.

Lothar Hotz, Dr. rer. nat. (*1960) is CEO and project leader at the Hamburger Informatik Technology-Center e.V. (HITeC) and the University of Hamburg. He researches mainly for topics in the field of Artificial Intelligence (i.e., machine learning, knowledge-based systems, constraints, planning, configuration systems, computer vision) and semantic-based Open Data portals.

3rd Conference on Production Systems and Logistics

PPC Task Plan Sourcing - Synchronization Of Procurement And Production. A Model-based Observation

Simon Hillnhagen[1], Selim-Hendrik Koller[1], Alexander Mütze[2], Peter Nyhuis[2], Matthias Schmidt[1]
[1]Leuphana University of Lüneburg, Institute of Product and Process Innovation (PPI), Lüneburg, Germany
[2]Leibniz University Hannover, Institute of Production Systems and Logistics (IFA), Hanover, Germany

Abstract

As companies continue to globalise, manufacturers face the challenge of strategically adjusting their vertical integration and restructuring production and supply chains. This leads manufacturers to increasingly pursue two strategies of restructuring. On the one hand, in the form of outsourcing value-adding activities, the focus is being placed on the core competencies of the company's own production. As a result, the vertical range of manufacture within the company is decreasing, while outsourcing is becoming more and more important. On the other hand, companies are also pursuing the strategy of least possible dependency to secure production through regional procurement of resources and expansion of the necessary competencies by means of increased vertical integration. In order to understand the consequences and effects of these changes at the level of production planning and control (PPC), a model-based view is necessary for an expanded understanding of the processual context of these changes. The PPC is the essential steering instance of production. It combines long-term tasks, e.g. Plan Sales or Roughly Plan Resources, with short-term tasks, e.g. Schedule Throughput or Plan Resources in detail. The main PPC task, Plan Sourcing, is an essential link with its tasks and procedures between the core processes of procurement and production in the company's internal supply chain.

In the context of this paper, the PPC main task Plan Sourcing is to be considered in a model-based manner, which focuses on the selection and connection of suppliers as well as the general view of the supplier management of manufacturers. For this purpose, the effect on the PPC and the production logistic objectives variables is presented by means of the consideration of the tasks and possible procedures for the fulfilment of these PPC tasks. Utilising collected findings, a process-related derivation for the synchronisation of the affected areas of procurement and production is presented.

Keywords

Procurement; Synchronization; Modelling; Production Planning and Control; PPC; Sourcing

1. Introduction

Globalisation has strongly changed the competition for manufacturing companies. In addition to rising cost pressure, companies are confronted with increasing individualisation requirements and shorter product life cycles [1,2]. In order to compete in this environment, it was a global trend, especially before the pandemic, for companies to focus on their core competencies in their production and outsource large sections of the value chain. During the pandemic, this trend has changed from a strongly global perspective to a locally oriented procurement in the area of particularly vulnerable supply chains. This is why it is currently called

DOI: https://doi.org/10.15488/12121

the "glocal / glocalisation" mega-trend, i.e. global and local procurement. [3,4] The reasons for outsourcing are that in many cases specialised suppliers can produce and deliver certain components much more cost-effectively than manufacturing these specific products in-house [5]. In addition, the suppliers, for their part, can drive forward the technical development of certain components in a more targeted manner and thus increase the quality of the final product [6]. For this reason, besides optimising the own production, procurement is becoming increasingly important, especially in complex supply chains [7].

In general, the task of procurement is to cover the demand for raw materials, semi-finished products and primary products that cannot be produced economically by the company's own production or cannot be produced at all [8]. Therefore, externally sourced items must reach production at the right time, in the right quantity and quality, in order to fulfil internal processes and specifications and to ensure the logistical efficiency of production [9]. Procurement is particularly significant for companies with multi-variant product portfolios and strong individualisation by customers, as they often carry out order-specific procurement for little-used articles and thus do not have any safety stocks in the warehouse. This can be observed more frequently especially with contract manufacturers (engineer-to-order (ETO) and make-to-order (MTO) manufacturers). In such cases, procurement is on the lead-time-critical path, which is why procurement processes and production processes should be synchronised as much as possible in order to improve the lead time and thus the response time of the company. [8,10] This paper will therefore analyse the interface between procurement and production in more detail. For this purpose, the process crossovers will be identified and examined in the wider context in relation to PPC.

2. Research Question and Research Methodology

In today's business structures, different departments often have overlapping subtasks where different objectives are always in the focus, leading to difficulties in coordination. In addition, the optimisation of corporate coordination is necessary, as this improves not only vertical but also horizontal cooperation with the supplier network. [2] The focus of this paper is the analysis of the interface between procurement and production. The contribution follows the research question: "How can procurement be optimised and synchronised based on control via production logistics objectives?”

In addition, the preconditions for this type of minimisation of interdepartmental conflicts are to be considered. Here, the focus is on procurement, since another interface exists here as a communication channel outside the company. Through a model-based approach, it is to be found out to what extent the production planning and control of the Hanoverian Supply Chain Model (HaSupMo) can make a target-oriented contribution in order to control logistics performance and logistics costs individually and to adapt them to the company-specific needs. Therefore, the main PPC task of “Plan Sourcing” of the HaSupMo will be presented later in chapter 4.1 and its characteristics will be shown as well as the resulting logistical objectives will be dealt with. In terms of the research question, a possible answer will be given with a qualitative approach based on the Hanoverian Supply Chain Model.

3. General Interface of Procurement and Production

The operational design of a company's supply chain involves overlapping cross-sectional tasks in which the respective decisions have a direct impact on objectives [8]. However, their influence on the logistical objectives in a company is not comprehensively known in industrial practice. This mutual, mostly unintentional, influence must be avoided through efficient and effective interface management. Joint coordination of the departments on necessary measures is indispensable to reduce internal barriers [11]. Departments such as procurement and production are essential constituents of every company. Therefore, these departments are required to engage in extensive exchange through transparent communication and coordination [12]. The reduction of interfaces is therefore an important approach to optimising internal processes [11,13].

3.1 Interface analysis

Companies have a large number of interfaces in connection with procurement, which can be found either externally, e.g. with suppliers, or internally, e.g. in relation to the supply chain between the warehouse and production. The primary objective of the company's internal supply chain should therefore be to optimally integrate procurement as a higher-level link in order to avoid supply interruptions in production. Especially for order-based producers with individual customer requirements, the lead time and adherence to delivery dates for the completion of the products is an important competitive factor [14]. To ensure that these sometimes sensitive process chains function in a resilient way, a comprehensive analysis is required as the basis for designing a target-oriented PPC. In order to systematically work out the characteristics of supplier activities as well as the necessary information flows, the value stream analysis and the SIPOC method (Supplier, Inputs, Process, Outputs, Customer), for example, which are anchored in the Six Sigma toolbox of methods, are valuable tools. These proven and clearly structured overview methods not only enable transparency of the essential interfaces in terms of content and process, but also show the data-driven input and output through the value creation process.

A closer look at processes and their interfaces always highlights the fact that the results usually relate to different levels of consideration and interact with each other. For the best corporate strategy is difficult to implement if workflows and processes are not clearly structured up to the operational level. On the other hand, corporate process analyses, such as value stream mapping, always start from the current production processes via coordination points such as planning and IT interfaces. This knowledge is essential in order to be able to make well-founded strategic management decisions.

3.2 Approach of the interfaces at four levels

The levels of the realization win for interfaces with production analyses, addressed in chapter 3.1, are to be pointed out now in this approach, at the example of production and procurement. In this approach, four levels can be identified, which build on each other and are interdependent (see figure 1). In addition, two procedural directions are presented. The analysis procedure from level 1 to level 4 (bottom to top) and the hierarchical control direction from level 4 to level 1 (top to bottom).

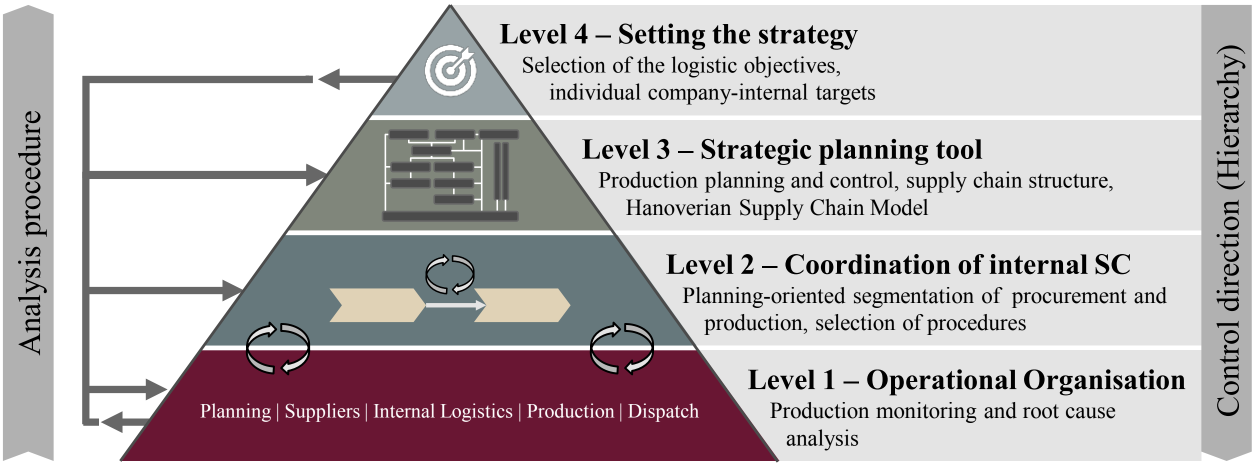


Figure 1: Four levels of interfaces (inspired by [15])

Level 1 forms the foundation for the production-related analysis of processes and interfaces. This level 1 "Operational Organization" is thus the starting point for the analysis direction. Level 4, on the other hand, is a strategic control instance and is to be regarded as the starting point for the direction of control (hierarchy).

After considering the processes at the production level, it is necessary to carry out segmentation at the product level and to categorize these products according to a defined value, for example. Based on this, a classification of components must be carried out in level 2. This classification of components can basically be divided into three categories. Parts that have to be procured externally due to the fact that there is a limited vertical range of manufacture. Parts that are manufactured in-house in any case due to the existing vertical range of manufacture. A third category consists of parts that can be sourced externally as well as manufactured in-house. On the one hand, products in the third category in particular bring the necessary flexibility to production in relation to the objective of capacity utilization. Depending on the allocation of these products to in-house production or external procurement, this is also accompanied by increased planning effort in procurement. This is because either raw materials or finished assemblies have to be procured.

In the next level 3, production planning and control is reviewed. Here, for example, the PPC of HaSupMo can be seen as a strategic tool. At this level, the planning stage is reached. Planning is based on data and assumptions of the past and also the future. The examination of the already configured PPC thus shows helpful clues for an alignment of the strategy of a company based on objectives. These clues are made visible and measurable by means of the PPC the objectives of the production logistics.

The fourth level now has the goal of defining a strategic orientation for the company. Here, the logistical objectives can be defined individually within the company. For example, companies with high quality requirements, such as aircraft manufacturers, should not necessarily pursue the objectives of short lead times or adherence to deadlines as a top priority. On the other hand, companies from the consumer goods segment pursue logistical objectives such as a high finished goods inventory and a short delivery time. This strategic objective orientation forms the framework and thus indicates the corresponding scope for decision-making for the departments with a focus on the interfaces.

In this way, using functioning process flows from Level 1, the classification of components from Level 2 and reduced interfaces by means of PPC tools from Level 3, production logistics objectives can be set by strategic business decisions. Based on the management's specification of the target values, the production programme can now be adjusted again and again by the three lower levels via iteration loops, and the type of procurement and production can be regularly reviewed and adjusted. At Level 1, processes and procedures in the operational organisation can also be changed and adapted to the procedures for optimising the specified objectives.

4. Integration of PPC

Production planning and control is an instrument that regulates the processes in the internal supply chain of manufacturing companies [16]. Their tasks include the entire planning and flow of production. From processing and clarifying orders to dispatching the finished products. It plans and controls orders taking into account the available resources and in compliance with overriding corporate objectives such as delivery time, punctuality and cost minimisation. [1,16] In addition to the planning and control of processes along the company's internal supply chain, production controlling, i.e. monitoring and intervening in the event of disruptions in the production process, is also part of the PPC's tasks [17]. Because PPC includes all areas of operational production management such as procurement, production, assembly and dispatch, it forms the organisational core of manufacturing companies [1,6,16].

Due to the complexity, scope and high importance of these interrelations, there are various approaches and concepts that model the tasks and processes within PPC. These range from more technical concepts, such as MRP-I and MRP-II logic, to various frameworks, such as the Aachen PPC model [16], and specialized modelling approaches for individual PPC tasks, such as the Manufacturing Control Model of Lödding [18]. An overview of the evolution of PPC can be found in [8], [18] and [19].

A model for describing the PPC, which integrates approaches from the Aachen PPC model and the modelling of LÖDDING, is presented in the framework "Hanoverian Supply Chain Model". This approach models for the first time the link between the eleven different main tasks of the PPC and the related objectives of the entire internal supply chain (see figure 2, right-hand side) [8]. The model shows the interrelationships between the PPC tasks and the control, regulating and objective variables. The target, planned and actual variables in the material flow as well as the effects on the logistical objectives in the individual core processes of the company's internal supply chain are placed in a causal relationship [8].

4.1 Plan Sourcing in HaSupMo

Focusing on managing the supplier network not only vertically but also horizontally, the share of external sourcing is becoming higher and higher, which is why the importance of procurement is steadily increasing [20]. In the PPC of HaSupMo, the central procurement tasks are controlled in the PPC main task "Plan Sourcing". As the main task of the PPC, it is the central interface between procurement and the company's internal production and therefore of great importance for the company's success [16]. The PPC main task "Plan Sourcing" is divided into five sub-tasks that cover the entire procurement process (see figure 2, left-hand side).

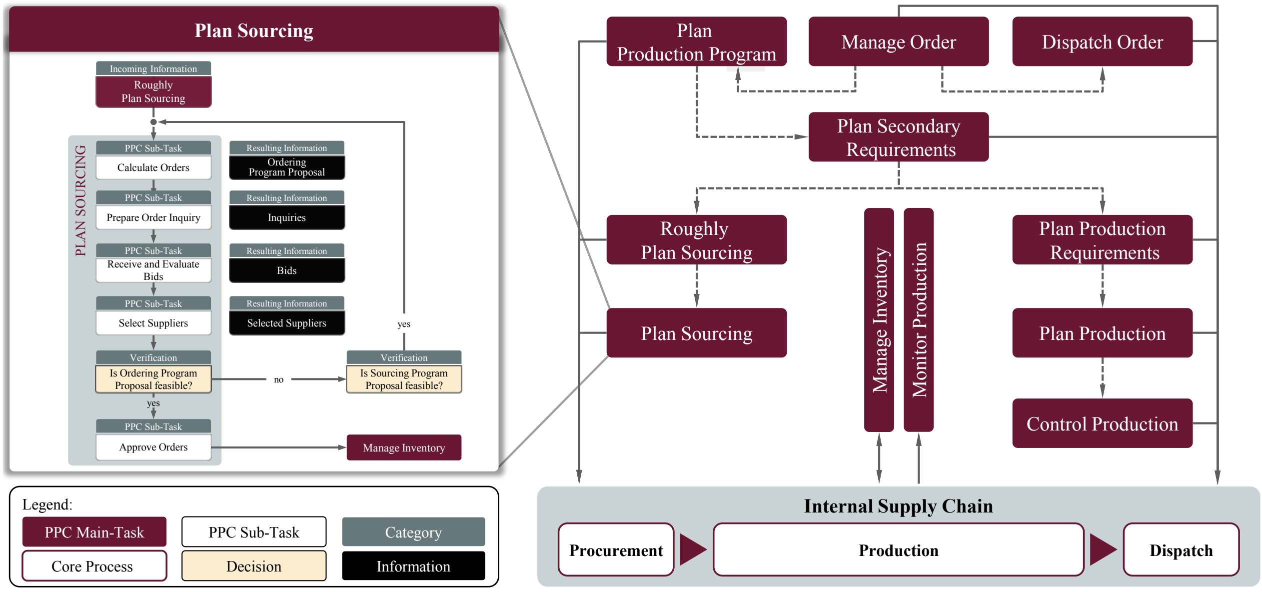


Figure 2: PPC-Main Task Plan Souring and the Hanoverian Supply Chain Model according to SCHMIDT AND SCHÄFERS [24] and SCHMIDT AND NYHUIS [8]

The first sub-task is "Calculate Orders". Here, an order programme proposal with the optimal order quantities and times for requirements to be procured is created from the determined primary and secondary requirements. From the order programme proposal, which contains the determined order quantities and times, in the following sub-task "Prepare Order Inquiry" concrete inquiries can then be made to the already selected suppliers. In the sub-task "Receive and Evaluate Bids", based on the inquiries, the suppliers can in turn create offers from which the company can subsequently select the most optimal offer according to the target premises. The "Select Suppliers" task is carried out from the analysis of the offers, taking into account the price, quality and logistical performance of the suppliers. [21,22,23] After the suppliers have been selected, the "Approve Orders" are placed. If no "Approve Orders" are placed, the system then checks whether the external procurement programme is feasible so that it can be carried out again from the "Calculate Orders". If the external procurement programme cannot be carried out, a message is sent to the PPC main task "Plan Production Requirements", which then checks the consequences for in-house production. As a result, the rescheduling of individual orders or in-house production of outstanding requirements can be initiated. [8]

4.2 Influence of logistical Objectives for Procurement

According to the PPC definition of HaSupMo, “Plan Sourcing” influences the objectives of inventory, due date Compliance and service level, or expressed the other way round, delivery delay, in the core process of procurement, the first core process in the company's internal supply chain [8]. As already mentioned, the task in procurement is to cover the needs that are not covered by the company's own production [1]. The inventory is formed from the stored articles [25]. Inventory results in inventory-induced costs, e.g. through capital commitment or inventory costs. Therefore, inventory should be kept as low as possible in order to keep inventory costs low from an economic point of view [1]. The service level is an important indicator of a company's logistics performance. It indicates what percentage of the demands from the inventory could be served in the right quantities and at the right time without a delivery delay or failure (see figure 3) [25,26].

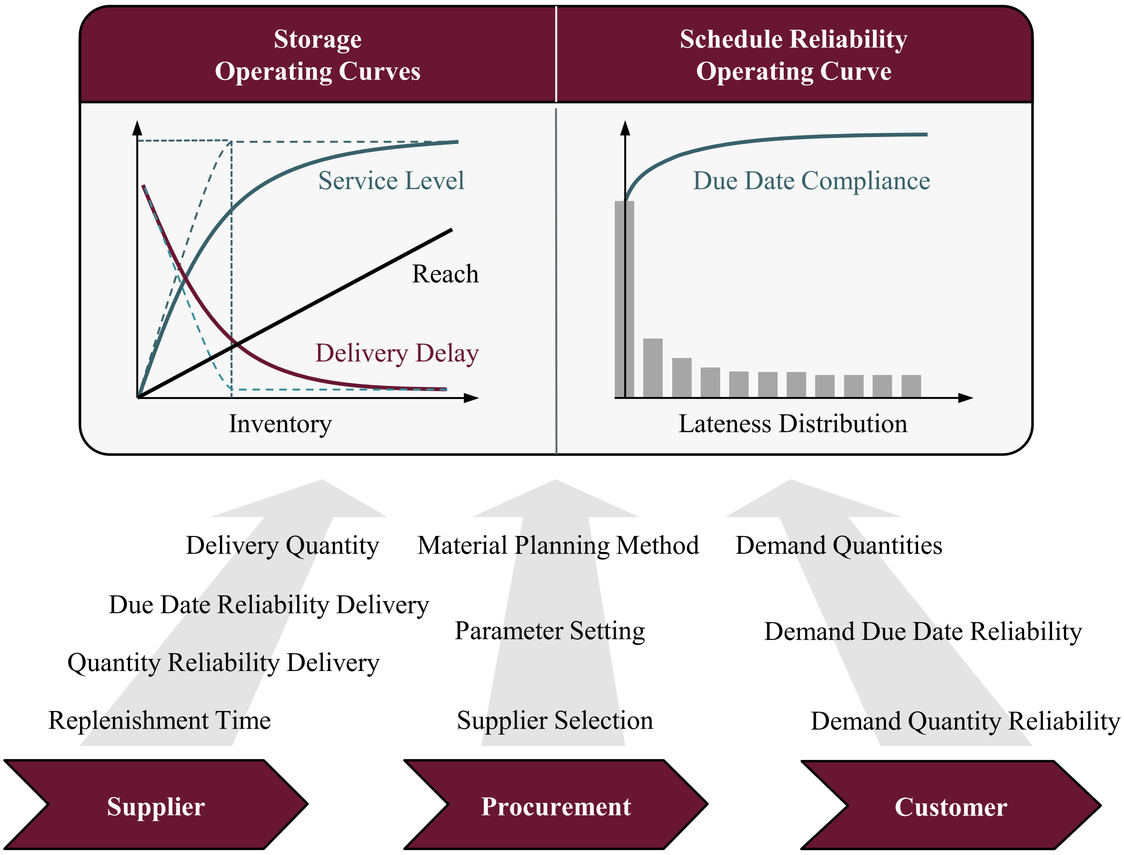


Figure 3: Influencing Factors of Storage and Schedule Reliability Operating Curves according to LUTZ [25] and NYHUIS AND WIENDAHL [27] and NYHUIS AND MAYER [28]

For this reason, the target is to achieve the highest possible level of service, as this demonstrates a high level of logistics performance and is synonymous with low delays in delivery from the warehouse compared to production [8]. The "due date compliance" results for order-specific procurement and shows the share of orders that are available in terms of quantity and on time [18]. Similar to the "service level", the aim is to achieve the highest possible "due date compliance". As the average inventory of stock items and directly procured items increases, the logistics performance in the form of the logistics objectives "service level" and delivery delay is positively influenced. In addition, this is accompanied by an increasing range, while at the same time logistics costs rise. This demonstrated influence of logistical objectives illustrates that a clear strategic orientation is of high importance for the configuration of the PPC. A clear positioning between the logistical objectives can only be achieved in this way.

4.3 Conflict of Objectives

In logistics target achievement, it is often necessary to position the company between logistics costs and logistics performance as shown in figure 4 [25]. In order to be able to achieve a high "service degree" in the procurement, the stock (high) is largely dependent on the access lot size and the performance of the supplier or the safety stock. By a transparent communication with the supplier and a purposeful production planning the performance can be besides substantially affected. If this is not the case, there is a risk of failures or delays if warehouse demands cannot be met. As a result, the objective "service level" cannot be sufficiently fulfilled. [8] The same situation exists for the objective "due date compliance". To ensure availability at the time of need and to avoid interruptions in production, safety times are included as buffers for procurement, which subsequently result in higher inventories and thus higher logistics costs. [1] In order to guarantee these conflicting objectives of high logistical efficiency and high logistics performance with low or optimal logistics costs, these differences must be minimised and optimally positioned to meet the company's internal needs.

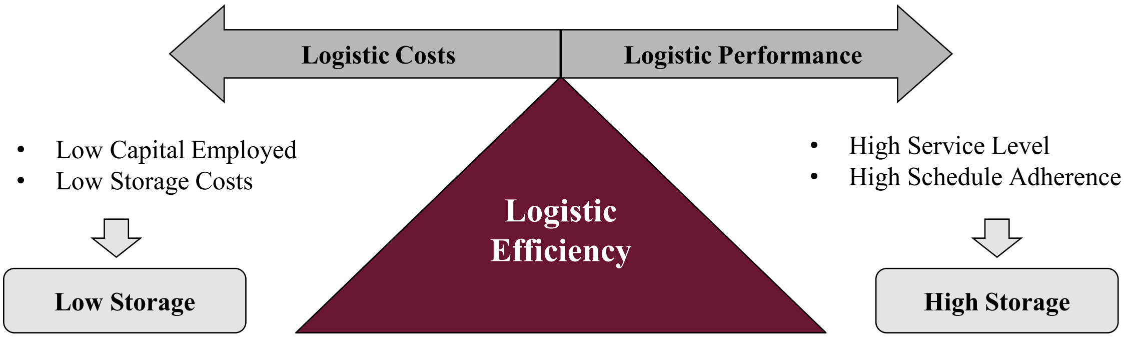


Figure 4: Logistical target system and conflicting objectives of procurement according to LUTZ [25]

For a holistic view of the conflicting objectives, an exact process-related localisation of the problem area is therefore important for a later synchronisation. The interaction of different areas of the company-internal supply chain often forms the area of tension, since procurement and production often pursue different interests and do not submit to the superordinate objectives through company specifications.

In order to obtain a standardised view of this interface between procurement and production, the production logistics objectives, as described in the HaSupMo, form a standard basis for evaluating and assessing these conflicting objectives. This is because direct conflicts are usually visible and can be made measurable. Indirect impacts, on the other hand, such as the influence on objectives, which is difficult to measure, can only be tracked via a strategic determination of defined objectives.

The approach presented in chapter 3.2 of using the PPC as a tool for checking the strategic alignment in controlling production for the production logistics objectives makes it possible to optimally match the required capacities with the order situation. Early information from production can be checked and adjusted with the interface to the supplier. With a lead time, determined capacities can be expanded or reduced within a certain framework at an early stage via external procurement. On the other hand, in-house production serves as a further balancing instrument that can cushion possible fluctuations. This type of general make-or-buy decision of dependent requirements makes it possible to achieve optimal values for production logistics goals and can thus be evaluated as an important mainstay. However, it must be taken into account that the framework for capacity adjustment of this kind is always tied to the flexibility of the products of the dependent requirements and the company's own vertical range of manufacture.

The order situation is often subject to predictable and unpredictable fluctuations in the course of a fiscal year. This is due to a variety of different factors. For a manufacturing company, however, capacity-adjusted utilization of the available resources in defined planning periods is necessary in order to avoid transferring fluctuations in incoming orders directly to production. As shown in chapter 3, the use of a PPC offers the possibility to detect any changes at an early stage and to react to them proactively. In order to optimally supply a production with the planned demand, not only a good and dense supplier network is necessary, but also a suitable configuration of the PPC, as this already provides a lot of information for the interface observation.

5. Conclusion and Outlook

In the context of this paper, a model-based analysis for the interface of procurement and production was conducted. Based on the challenges in the production-related environment, the necessary prerequisites and structures were pointed out that are required for a sensible integration and configuration of a PPC.

The approach to the interface analysis presented in chapter 3.2 points out in the first step that interfaces from the operational level up to the strategy definition should be regarded holistic. This approach offers the advantage of being able to control the defined corporate strategies hierarchically in the further development. Especially the question of the make-or-buy decision for affected components significantly influenced the capacities in defined time periods. However, these affected components have the positive side effect of being able to compensate for fluctuations that occur by means of targeted control of the PPC deficits or capacities. Based on the HaSupMo, the interface of the dependent requirements planning was localised as the interface between production and procurement. Depending on the allocation of products, the main PPC task of Plan Sourcing presented in chapter 4.1 is significantly influenced. These effects also represent an influence on the production logistics objectives, which serve as a control option.

Nevertheless, the results presented are characterised by basic theory that still needs to be transformed for practical use in the future. Further research activities are to be carried out in the validation of the presented results in the production environment. Despite this, it can be stated that the entrepreneurial focus should be on a sophisticated PPC and its influence on logistical objectives to be able to react successfully and proactively to dynamic market movements.

Acknowledgements

The project “Systematic analysis of the effect of production planning and control processes on logistical objectives” leading to this paper is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 434659386.

References

[1] Wiendahl, H.-P., Wiendahl, H.-H., 2019. Betriebsorganisation für Ingenieure, 9., vollständig überarbeitete Auflage ed. Carl Hanser Verlag München, München, 425 pp.

[2] Handfield, R., Straube, F., Pfohl, H.-C., Wieland, A., 2013. Embracing global logistics complexity to drive market advantage. DVV Media Group, Hamburg, 81 pp.

[3] Vanchan, V., 2021. Global pandemic disruptions, reconfiguration and glocalization of production networks, in: Bryson, J.R., Andres, L., Ersoy, A., Reardon, L. (Eds.), Living with pandemics. Places, people and policy. Edward Elgar Publishing, Northampton, pp. 195–201.

[4] Helmold, M., Terry, B., 2021. Global Supply Chain and Logistics, in: Helmold, M., Terry, B. (Eds.), Operations and Supply Management 4.0. Industry Insights, Case Studies and Best Practices, 1st ed. 2021 ed. Springer International Publishing; Imprint Springer, Cham, pp. 107–114.

[5] Koether, R., 2018. Distributionslogistik: Effiziente Absicherung der Lieferfähigkeit, 3., aktualisierte und erweiterte Auflage ed. Springer Fachmedien, Wiesbaden, 252 pp.

[6] Vahrenkamp, R., Siepermann, C., 2008. Produktionsmanagement, 6., überarbeitete Auflage ed. Oldenbourg Wissenschaftsverlag, München.

[7] Wannenwetsch, H., 2014. Integrierte Materialwirtschaft, Logistik und Beschaffung, 5.th ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 796 pp.

[8] Schmidt, M., Nyhuis, P., 2021. Produktionsplanung und -steuerung im Hannoveraner Lieferkettenmodell. Springer Berlin Heidelberg, Berlin, Heidelberg, 225 pp.

[9] Luczak, H., Eversheim, W. (Eds.), 2001. Produktionsplanung und -steuerung: Grundlagen, Gestaltung und Konzepte, 2., korrigierte Auflage 1999, Nachdruck ed. Springer Berlin Heidelberg, Berlin, Heidelberg, s.l., 773 pp.

[10] Rennemann, T., 2007. Logistische Lieferantenauswahl in globalen Produktionsnetzwerken. DUV, Wiesbaden.

[11] Albers, S., Gassmann, O. (Eds.), 2005. Handbuch Technologie- und Innovationsmanagement: Strategie - Umsetzung - Controlling. Gabler Verlag, Wiesbaden, s.l., 849 pp.

[12] Meckel, M. (Ed.), 2008. Unternehmenskommunikation: Kommunikationsmanagement aus Sicht der Unternehmensführung, 2., überarb. und erw. Aufl. ed. Gabler, Wiesbaden, 556 pp.

[13] Specht, G., Beckmann, C., Amelingmeyer, J., 2002. F&E-Management: Kompetenz im Innovationsmanagement, 2., überarb. und erw. Aufl. ed. Schäffer-Poeschel, Stuttgart, 570 pp.

[14] Nebl, T., 2011. Produktionswirtschaft, 7., ed. Oldenbourg Wissenschaftsverlag, München, 976 pp.

[15] Mütze, A., Lucht, T., Nyhuis, P., 2022. Logistics-Oriented Production Configuration Using the Example of MRO Service Providers. IEEE Access 10, 20328–20344.

[16] Schuh, G., Stich, V., 2012. Produktionsplanung und -steuerung 1: Grundlagen der PPS, 4., überarbeitete Auflage ed. Springer, Berlin, Heidelberg, 491 pp.

[17] Zäpfel, G., 2001. Grundzüge des Produktions- und Logistikmanagement, 2., unwesentlich veränd. Aufl. ed. Oldenbourg, München, 289 pp.

[18] Lödding, H., 2013. Handbook of Manufacturing Control: Fundamentals, description, configuration. Springer Berlin Heidelberg, 577 pp.

[19] Olhager, J., 2013. Evolution of operations planning and control: from production to supply chains. International Journal of Production Research 51 (23-24), 6836–6843.

[20] Lasch, R., 2019. Strategisches und operatives Logistikmanagement: Beschaffung, 2., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden, 345 pp.

[21] Wannenwetsch, H., 2021. Integrierte Materialwirtschaft, Logistik, Beschaffung und Produktion: Supply Chain im Zeitalter der Digitalisierung, 6. Aufl. 2021 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 909 pp.

[22] Irlinger, W., 2012. Kausalmodelle zur Lieferantenbewertung. Zugl.: Passau, Univ., Diss., 2011, 1. Aufl. ed. Gabler, Wiesbaden, 179 pp.

[23] Janker, C. (Ed.), 2008. Multivariate Lieferantenbewertung: Empirisch gestütze Konzeption eines anforderungsgerechten Bewertungssystems. Zugl.: Dresden, Techn. Univ., Diss., 2004, 2., aktualisierte und erw. Aufl. ed. Gabler, Wiesbaden, 411 pp.

[24] Schmidt, M., Schäfers, P., 2017. The Hanoverian Supply Chain Model: modelling the impact of production planning and control on a supply chain's logistic objectives. Prod. Eng. Res. Devel. 11 (4-5), 487–493.

[25] Lutz, S., 2002. Kennliniengestütztes Lagermanagement. Zugl.: Hannover, Univ., Diss., 2002, Als Ms. gedr ed. VDI-Verl., Düsseldorf, 159 pp.

[26] Schönsleben, P., 2016. Integral Logistics Management: Operations and Supply Chain Management Within and Across Companies, Fifth Edition, 5th ed. ed. CRC Press, Boca Raton, 844 pp.

[27] Nyhuis, P., Wiendahl, H.-P., 2009. Fundamentals of Production Logistics: Theory, Tools and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, 320 pp.

[28] Nyhuis, P., Mayer, J., 2017. Logistic modelling of lateness distributions in inventory systems. Prod. Eng. Res. Devel. 11 (3), 343–355.

Biography

Simon Hillnhagen (*1989) studied industrial engineering at the Leuphana University of Lüneburg and has been working as a research assistant at the Institute for Product and Process Innovation (PPI) since 2020 in the field of production management with a focus on production planning and control.

Selim-Hendrik Koller (*1998) studied Business Administration at Leuphana University of Lüneburg. Since 2021 he is studying European Management at the Eberhard-Karls-University of Tübingen and the EM Strasbourg Business School.

Alexander Mütze (*1994) studied industrial engineering at Leibniz University Hannover and has been working as a research assistant at the Institute of Production Systems and Logistics (IFA) since 2018 in the field of production management with a focus on production planning and control.

Peter Nyhuis (*1957) studied mechanical engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After obtaining his Dr.-Ing. doctorate, he was habilitated before working as an executive in the field of supply chain management in the electronics and mechanical engineering industry. He has been head of IFA since 2003.

Matthias Schmidt (*1978) studied industrial engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he became head of Research and Industry of the IFA and received his habilitation. Since 2018, he holds the chair of production management at the Institute for Product and Process Innovation (PPI) at the Leuphana University of Lüneburg. In addition, he became the head of the PPI in 2019.

3rd Conference on Production Systems and Logistics

Typification Of Incorrect Event Data In Supply Chain Event Management

Jokim Janßen[1], Günther Schuh[1], Volker Stich[1]
[1] *Institute for Industrial Management (FIR) at RWTH Aachen, Aachen, Germany*

Abstract

Due to shorter product life cycles and the increasing internationalization of competition, companies are confronted with increasing complexity in supply chain management. Event-based systems are used to reduce this complexity and to support employees' decisions. Such event-based systems include tracking & tracing systems on the one hand and supply chain event management on the other. Tracking & tracing systems only have the functions of monitoring and reporting deviations, whereas supply chain event management systems also function as simulation, control, and measurement. The central element connecting these systems is the event. It forms the information basis for mapping and matching the process sequences in the event-based systems. The events received from the supply chain partner form the basis for all downstream steps and must, therefore, contain the correct data. Since the data quality is insufficient in numerous use cases and incorrect data in supply chain event management is not considered in the literature, this paper deals with the description and typification of incorrect event data. Based on a systematic literature review, typical sources of errors in the acquisition and transmission of event data are discussed. The results are then applied to event data so that a typification of incorrect event types is possible. The results help to significantly improve event-based systems for use in practice by preventing incorrect reactions through the detection of incorrect event data.

Keywords

CPSL; Supply Chain Event Management; Incorrect Data; Event Data; EPCIS; Typification.

1. Introduction

The supply chain management sector is becoming more dynamic and complex as a result of rising globalization. This makes the framework conditions more unstable, which raises the probability of unplanned processes [1]. This results in an expansion of the effort required for process control as well as an increase in the number of interfaces between companies and the number of processes to be controlled [2]. These conditions as well as an increase in networking lead to an ever-greater amount of data being absorbed by companies, from which the relevant data has to be extracted. The oversupply of information needs to be reduced to information that deviates and requires the attention of a decision-maker. This can be achieved through event-based systems working according to the concept of management-by-exception [3]. These systems include, for example, tracking & tracing systems and supply chain event management. The information technology basis for these systems is formed by events, which contain the data arising from a planned or unplanned event in a standardized format [4]. For the information to be used reliably in the following processes and for decisions to be made correctly, the information must first be available and the

DOI: https://doi.org/10.15488/12183

data it contains must be of high data quality [5]. If this is not the case, it can lead to serious consequences, as the following example shows: A supplier ships the product to the manufacturer, but for various reasons, the corresponding event message is not transmitted to the manufacturer. According to the event-based systems, the manufacturer starts to reschedule his production to avoid a production stop. A day later, the event is transmitted with a delay, so the manufacturer again reschedules his production. This example shows that event data that does not correspond to reality can lead to significant problems and avoidable additional costs. This can also be seen in practice. This can also be seen in practice, where the data exchange with event data does not work properly nowadays. The desired data is often not available, the interfaces are not sufficiently defined or the quality of the data is insufficient [6]. In particular, the aspect of data quality has so far been completely ignored in the description of events [7].

2. Basic Concepts Regarding Event-Based Systems

For the typification of incorrect event data, it is essential to have an overall understanding of the relevant terms. Concerning the term event, it is important to understand the definition of the term and the data structure of event standards (cf. section 2.1). To be able to comprehend the analysis of incorrect event data, the concept of supply chain event management and its difference from other event-based systems must also be considered (cf. section 2.2).

2.1 Events

The term “event” regarding event-based systems is not uniformly defined in the literature. On the one hand, events are described as occurring activities in the real world or a computer system [8]. According to Bensel et al., the term can be described as the associated data object for the occurrence of a state with essential significance for logistical processes [9]. On the other hand, Heusler et al., for example, expound that an event can only be understood as a deviation from a planned state [10]. The definitions in the literature can be sorted into the categories 'event in the sense of a status report' and 'event in the sense of a deviation' – depending on the main statement. As this paper focuses on the relevant data objects and the associated standards of events, the understanding of ‘event in the sense of a status report’ is being followed.

Within the scope of the exchange of events, different standards exist, which ensure that the sender and receiver use compatible formats so that the events can be read without any loss in the receiver system [11]. Through literature research, Konovalenko & Ludwig were able to identify three common standards: Tracefish, TraceCore XML, and EPCIS; the latter being the most widely used. One reason for this is that this standard is subject to fewer restrictions than the other two. Therefore, it can be used universally in various industries. In contrast, TraceCore XML was designed for data exchange in the food industry with a focus on food traceability; and Tracefish's standard is specific to the fish industry. [7] In addition to the standards mentioned above, there is also the Health Level 7 standard, which is only used in the health sector [12]. As the overview of the various event standards has shown, only the EPCIS standard is independent of the industry. Moreover, it is the most widely used standard. Thus, only the EPCIS standard will be considered in this paper.

According to the EPCIS standard (version 1.2), events always have a basic structure: each of the recorded events contains information from the four dimensions of what, when, where, and why using specific data elements [13]. The ‘dimension what’ specifies which objects or which object classes are involved in the event. For example, the uniquely named objects or object classes are listed in an *epcList* and usually described by an EPC in the form of a uniform resource identifier. The date and time when the EPCIS event was created (*eventTime*) and recorded in an EPCIS repository (*recordTime*) are stored in the ‘dimension when’. Additionally, the time zone of the location where the event was recorded (*eventTimeZoneOffset*) is recorded. The ‘dimension where’ specifies the exact capture point where the EPCIS event was generated (*readPoint*). In addition, the location of the business (*businessLocation*) where the object is now located is recorded. The ‘dimension why’ specifies the reason why the EPCIS event was created. The business process step (*businessStep*) in which the event was generated is recorded. Furthermore, the status (*disposition*) of the object is described. In addition, the *businessTransactionList* is used to assign a business transaction to the event. The *sourceList* and *destinationList* are used to provide additional business context if an EPCIS event is part of a business transfer of ownership, jurisdiction, or custody. The entire data is stored in an XML structure with a defined syntax, the vocabulary of which is specified by the CBV. [14] The four dimensions can be used to describe the content of each event that occurs in a physical or virtual object. All events can be divided into different types of EPCIS events. The ObjectEvent is the simplest and most commonly used event. It refers to a single or several objects and is responsible for the pure observation of these. Within the scope of the AggretationEvent, the physical merging or separation of one or more objects can be recorded. This type is the second most widespread standard and, together with the ObjectEvent, covers the majority of events that occur. Another type of event is the TransformationEvent. It is used when input objects are partially or completely consumed in the creation of output objects so that some or all of the input objects have contributed to each of the output objects. A TransactionEvent occurs when one or more objects are linked to or unlinked from one or more business transactions. [15]

2.2 Supply Chain Event Management

According to Stölzle et al., supply chain event management can be defined as follows: Supply chain event management (SCEM) is a tool for controlling logistical processes that enables the timely reaction to critical exceptional events in supply chains [16]. Accordingly, SCEM, like "tracking & tracing", is one of the event-based approaches, whereby SCEM is understood as a further development of track & trace (cf. Figure 1) as it provides the data for system-controlled decision support [17].

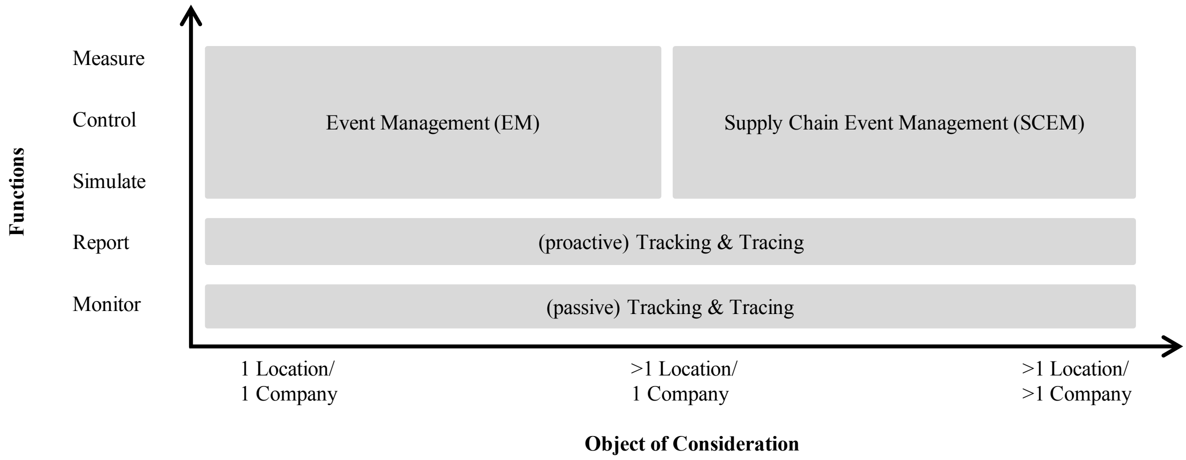


Figure 1:Classification of event-based systems (own figure)

The term SCEM system, in turn, refers to an information system or the set of information systems that enable the fulfillment of functionalities according to their function [18]. Concerning this fulfillment, the majority of the literature defines the following five core functions [10,7,18]:

- Monitor: The core function monitoring includes, on the one hand, the recording of the actual state and, on the other hand, the comparison with the target state including the defined tolerance window. Based on the comparison, an assessment of the deviation takes place afterward.
- Report: If a critical deviation from the plan is detected during monitoring, the reporting function takes over the real-time transmission of the information to the respective decision-making authority, so that it can actively intervene in the system and reduce the risk of major disruptions.

- Simulate: Following the registration of an event and notification, simulate checks and evaluates possible options for responding to the event, which serves as decision support.
- Control: The core element of controlling consists of selecting and implementing the promising alternative to correct the target-actual deviation in the best possible way.
- Measure: In measuring, the events are extracted from the individual actual runs. Based on a large number of observed runs, performance indicators are determined, which can be used either for evaluations of the supply chain processes or as input parameters for the core function monitoring.

Based on the described core functions, the SCEM can be classified in reference models, such as the supply chain planning matrix, between the areas of supply chain execution (short-term) and supply chain planning (medium-term) along the complete process chain [19]. The complete process chain explicitly means that this does not only concern supply chain processes but also all internal and external processes of order fulfillment. The reason for this classification is, on the one hand, based on the short-term reactions in the case of serious deviations and, on the other hand, on the additional possibility of abstracting actual processes, which leads to medium-term process improvements. Conceptually, the SCEM can be assigned not only in reference models but also in the following theoretical approaches: 'cybernetic control loop' and 'management by exception' [20,18].

3. Methodology

This paper addresses the following research question for the analysis of incorrect event data.: How do incorrect event data occur in supply chain event management and how can they be typified or described?

To answer this research question, a systematic literature review has been conducted. Due to the scientific recognition and the professional proximity, it is oriented towards the procedures mentioned by Webster & Watson and Levy & Ellis [21,22]. The factors used to define the study area can be summarised as follows:

- Inclusion criteria: language (German and English), availability (full text), document type (journal articles, monographs, collected works), period (2011-2021)
- Databases: ScienceDirect, IEEE Xplore, Google Scholar, RWTH Aachen University Library Catalogue
- Search terms: 17 search terms for data acquisition, 12 search terms for data transmission, 12 search terms for concrete forms of incorrect data/search terms are based on Boolean combinations from different keywords (study area (“supply chain”, “event management”, “EPCIS”, etc.), data (“transaction data”, “feedback data”, “event data”, etc.), errors (“deficient”, “inaccurate”, “inconsistency”, etc.), data acquisition (“data acquisition”, “data collection”, “production data acquisition”, etc.), data transmission (“data sharing”, “ETL”, “data exchange”, etc.), incorrect occurrences (“types”, “characteristics”, “error detection”, etc.))
- Search strategy: first search run (abstract review, full text review) / second search run (forward-backward search) / third search run (forward-backward search)

Regarding the research question, the search is split into three aspects. First, the modes of operation and sources of error in data acquisition and transmission are considered to find causes for the emergence of incorrect event data. Then, the concrete forms of incorrect data are being searched for. In the first search run on the topic of data collection, 31 relevant sources were found. In the second run, an additional 29 contributions were found, and in the third run, another 14 were discovered. In the search in the subject area of data transmission, a total of 18 relevant sources were found in the first run. In the second run, a further eight sources were found. The third run added two contributions. The search for manifestations of erroneous event data resulted in four new sources in the first run, eight sources in the second run, and no sources in the third run. In this way, a total of 114 contributions were found that are related to the research topic of this paper.

4. Research Results & Typification

As a result of the systematic literature research, typical sources of error in data acquisition and data transmission could be identified. In addition, approaches were considered that describe concrete forms of occurrence of incorrect event data.

4.1 Sources of Data Acquisition Errors

In practice, a wide range of data acquisition methods is used. To limit the identification of data acquisition error sources in this paper to those most relevant to practice, the following basic assumption was made: Data acquisition methods that are very often used in practice produce more data acquisition errors in absolute terms than those that are used less frequently. Based on this assumption, the sources collected in the literature review were also analyzed concerning the question of how frequently a data acquisition method was mentioned.

The six data acquisition technologies that are cited the most in the 114 sources surveyed are the following: RFID (62 sources, representing about 54%), 1D code (29%), manual data acquisition (25%), semi-automated operational data acquisition (18%), 2D code (10%), and RTLS (9%). Under the assumption made earlier, it can now be assumed that RFID technology causes the most data capture errors in practice, followed by 1D code (e.g. barcodes), manual data acquisition, etc.

Examples of specific errors in data acquisition can be easily described using the barcode or manual data acquisition. Typical sources of errors in barcode scanning are deterioration of the readability or damage to the code during transport due to scratches, dirt, or moisture [23]. In addition, the reader may be defective or incorrectly aligned during the reading process [24]. These sources of error can result in an event not being captured during acquisition (missing data acquisition) or not being passed to downstream systems in real-time (delayed data acquisition). The manual form of data acquisition is to be classified as particularly error-prone due to the high dependence on the human work factor. Thus, the correctness of data acquisition depends on the attention and ability of the person responsible [25]. This means, for example, that a lack of attention can lead to numerical errors during data entry. This in turn leads to incorrect data acquisition. Furthermore, it is possible that an object is captured although it should not be captured (unnecessary data acquisition) or is already captured (duplicate data acquisition).

Even when considering all of the different technologies, all of the data acquisition results related to event data can be grouped into six general categories:

- Correct Data Acquisition: Exactly the data that should be captured has been captured and is now correctly available.
- Missing Data Acquisition: The data that should have been captured was not captured and is therefore not present.
- Unnecessary Data Acquisition: Data that should not have been captured was captured anyway and is now unintentionally present.
- Duplicate Data Acquisition: Data was captured multiple times and is now redundant.
- Incorrect Data Acquisition: Data has been collected but is incorrect, inconsistent, or incomplete.
- Delayed Data Acquisition: Data was captured correctly, but is not available in real-time.

4.2 Sources of Data Transmission Errors

Data transmission can be divided into the steps of data integration and data exchange, whereby data integration is of higher importance in the context of this paper [26]. First of all, it should be noted that no relevant sources of error in connection with data exchange in the EPCIS standard could be identified during the literature research. One reason for this may be the high degree of standardization and automation of the standard and the associated lower probability of errors occurring.

Data integration can be divided into the substeps extraction, transformation, and loading (ETL) [27]. Failures in the ETL process can manifest themselves, on the one hand, in the fact that data that has already been fed in incorrectly is not recognized and corrected as intended and, on the other hand, in the fact that further errors are caused in the course of the process [28]. The causes for this can be found both in the development phase and in the operational phase of the ETL process and, similar to data acquisition, are based on both technical and human error [28]. Specific sources of failure in the development phase of the process include inadequate requirements definition [27], lack of testing of the process [29,30], and lack of continuous maintenance and adaptation of the process to changing user requirements [31]. Due to the lack of testing with real data, it is not possible, for example, to determine whether incorrect or redundant data is detected and cleaned up (incorrect data transmission). During the operational phase, errors may occur in the first two phases of the ETL process due to incorrect and/or inappropriate extraction and transformation rules [31–33]. For example, an incorrect assignment of data fields (schema mapping) can result in new incorrect data records (error-producing data transmission). In addition, technical sources of error, such as network errors or problems with the data carriers, can occur in the phase of loading the data into the repository [32]. In addition to creating new errors, this can also lead to delays so that the data is not provided in real-time (delayed data transmission)

When considering the sources of error in data transmission, it turned out that in general they can be divided into four possible categories, which are presented below:

- Optimal Data Transmission: Error-prone and redundant data is detected and cleaned up.
- Incorrect Data Transmission: Error-prone and redundant data is detected but not cleaned or not detected and not cleaned.
- Error-producing Data Transmission: Error-prone and redundant data is only produced in the course of data transmission.
- Delayed Data Transmission: Data is not provided in real-time due to delays in the process.

4.3 Specific Forms of Incorrect Data

Based on the typical sources of errors in data acquisition and transmission, it is now possible to concatenate manifestations of incorrect event data. For the categorization of the forms of occurrence of operational feedback data, there are some isolated approaches in the literature. These approaches subdivide the forms of occurrence based on data acquisition errors [34], data quality characteristics [35], data sources [29], or data attributes [36].

The occurrences of incorrect data found in the literature can be abstracted and summarized. It can be seen that errors in event data in the context of the SCEM can affect several data components: identification number (e.g., tag ID, item number), time information (e.g. date, time of entry), location information (e.g. reader ID, capture location), total data record. The type of error can vary between the six characteristics missing, wrong, redundant, inconsistent, unnecessary, and outdated. For example, incorrect event data can consist of missing identification numbers, redundant time information, or outdated location information.

4.4 Typification

Regarding the typification based on the forms of occurrence of incorrect event data, the causes of the emergence of incorrect event data and their occurrences have to be put into context and abstracted. Based on this, the occurrence of incorrect event data can be typified. For this purpose, the error sources of data acquisition and transmission are assigned to the occurrences of incorrect data (cf. Figure 2).

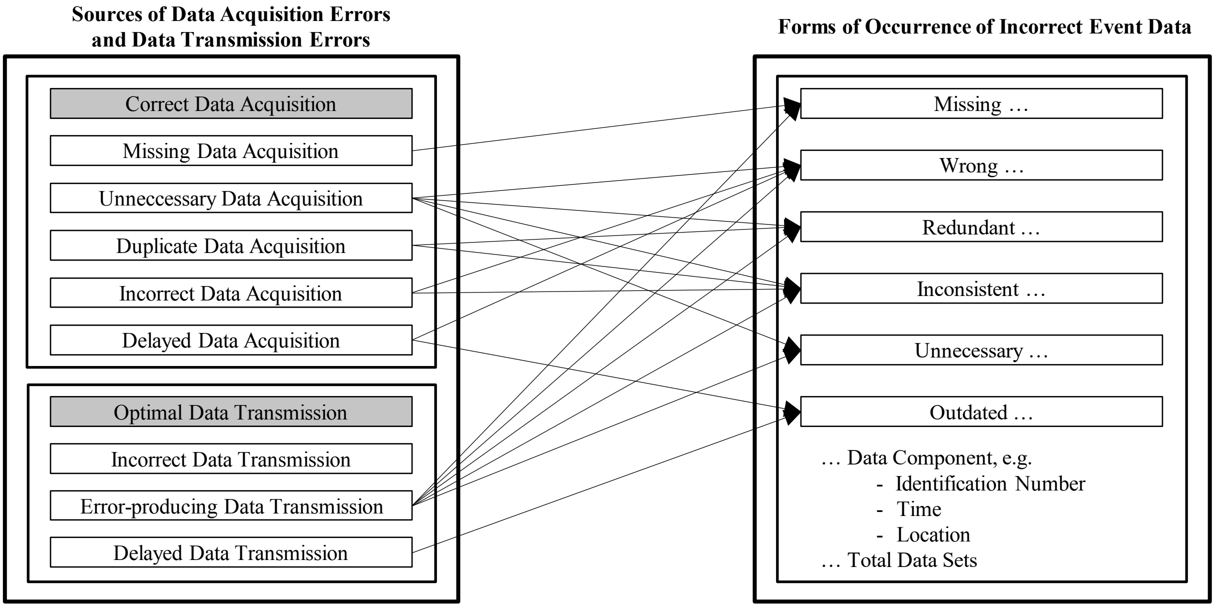


Figure 2: Framework for contextualizing and describing incorrect event data in SCEM (own figure)

The categories defined in chapters 4.1 and 4.2 regarding data acquisition and transmission are shown on the left side of the figure. Arrow connections are used to assign these to their possible occurrences in the context of incorrect event data (cf. chapter 4.3). For example, missing data acquisition can refer either to individual contents of a data set that are not recorded or to the entire data set. Specifically, it can therefore be associated with missing data components or missing overall data sets. If incorrect data is either not recognized and thereupon not cleaned or recognized and nevertheless not cleaned, then the incorrect data transmission would be present. It does not establish occurrences of incorrect event data, but it is also unable to correct such occurrences.

Using this framework, the incorrect event data can be typed and contextualized. Since the data components mentioned in chapter 4.3 form comparable categories as the dimensions of the EPCIS standard [15], specific incorrect event data can be derived. For this purpose, the four event dimensions are considered separately. This is briefly explained for each dimension using a specific example:

The ‘dimension what’ contains information about the physical and digital objects involved in an event [13]. A parallel can be drawn between the identification keys of the EPC and the data component of the identification number. When placed in the framework for contextualizing and describing incorrect event data in SCEM, the EPC identification key would, thus, be associated with the same occurrences and causes as the identification number. Therefore, an example of an incorrect occurrence of an EPCIS event could be a record that contains an incorrect SGTIN number (cf. Table 1).

Table 1: Example of an incorrect occurrence form of the ‘dimension what’

Dimension	Data Element	Incorrect Event Data	Correct Event Data
What	epcList	GTIN 106141411234569 Serial 12345	GTIN 106141411234569 Serial 12346

The ‘dimension when’ of EPCIS events includes the three elements of *eventTime, eventTimeZoneOffset,* and *recordTime* [14]. The contents of the dimension can be compared to the when data component in Figure 2. Based on a transfer of the findings to the ‘dimension when’, the occurrence of an outdated specification of the *eventTime* in the EPCIS standard could be mentioned (cf. Table 2). This could occur, for example, if an event is not recorded until sometime after its actual occurrence due to a time-delayed data collection.

Table 2: Example of an incorrect occurrence form of the ‘dimension when’

Dimension	Data Element	Incorrect Event Data	Correct Event Data
When	Event Time	Sep 23, 2012, at 10:12 am UTC	Sep 23, 2012, at 09:59 am UTC

The *readPoint* and *businessLocation* of an object are recorded in the EPCIS standard based on the ‘dimension where’. A possible occurrence of an incorrect EPCIS event could be, for example, an inconsistent specification of the Read Point if it contradicts the time zone specified in the When dimension. Table 3 shows a *readPoint* in Germany and at the same time an *eventTimeZoneOffset* that does not correspond to the German time zone.

Table 3: Example of an incorrect occurrence form of the ‘dimension where’

Dimension	Data Element	Incorrect Event Data	Correct Event Data
Where	ReadPoint	50° 46' 31.244" N 6° 5' 1.992" E	28° 0′ 44″ N, 28° 58′ 34″ E
When	*eventTime ZoneOffset*	+03:00 (UTC)	

The ‘dimension why’ can contain information about the *businessStep* and the *disposition* as well as a *businessTransactionList*, a *sourceList,* and a *destinationList*. For example, the occurrence of a record could be derived with a misstatement of *disposition* that indicates an object is stolen but is in transit between two trading partners (Table 9).

Table 4: Example of an incorrect occurrence form of the ‘dimension what’

Dimension	Data Element	Incorrect Event Data	Correct Event Data
Why	disposition	stolen	in_transit

5. **Conclusion & Outlook**

The concept of SCEM promises companies optimized, faster decision-making, enabled with the help of proactive notifications of relevant events within SC. In practice, this should be seen as an opportunity, especially against the backdrop of increasing customer demands and uncertain market conditions, to exploit the potential of SCEM and improve its competitive position in the long term. Companies are dependent on a solid database that provides the relevant data on time and with sufficient data quality.

This paper creates added value for the conceptual consideration of supply chain event management as well as its use in practice through the scientific consideration of incorrect event data. For the further development of event-based systems, it is indispensable to transfer the results concerning the potential sources of errors into the corresponding occurrences for event data. Through systematic consideration, the foundation for a generally valid typification of incorrect event data based on their form of occurrences has been laid. In the further, for example, automatic filtering of incorrect event data can be developed, which is before the actual core functions simulate, control and measure. This could exclude the automatic incorrect reaction based on incorrect event data (cf. example in the introduction).

When considering this field of research, there is a need for further research. For example, further consideration of possible causes of the emergence of incorrect event data could include other stages of the life cycle of the data, in addition to the steps of data acquisition and transmission.

References

[1] Bretzke, W.-R., 2020. Logistische Netzwerke, 4. Auflage ed. Springer Vieweg, Berlin, 1Online-Ressource (XIX, 541 Seiten).

[2] Wegner, U., Wegner, K., 2017. Einführung in das Logistik-Management: Prozesse - Strukturen - Anwendungen, 3. aktualis. u. erw. Auflage ed. Springer Gabler, Wiesbaden, 323 pp.

[3] Jeuschede, G., 1994. Grundlagen der Führung: Führungsprozeß, Führungskreis, Führungsfunktion, Führungskonzeptionen - Management by Objectives - Management by Exception - Management by Delegation - Führen nach dem Regelkreismodell, Führungsstil, 1. Auflage ed. Springer Fachmedien, Wiesbaden, 74 pp.

[4] Nissen, V., 2002. Supply Chain Event Management als Beispiel für Electronic Business in der Logistik, in: Gabriel, R., Hoppe, U. (Eds.), Electronic Business. Theoretische Aspekte und Anwendungen in der betrieblichen Praxis. Physica, Heidelberg, pp. 429–445.

[5] Nyhuis, P., Schmidt, M., Hübner, M., 2017. Transparenz durch Datenverfügbarkeit als Enabler für eine leistungsfähigere PPS, in: Reinhart, G. (Ed.), Handbuch Industrie 4.0. Geschäftsmodelle, Prozesse, Technik. Hanser, München, pp. 33–34.

[6] Kersten, W., Seiter, M., See, B. von, Hackius, N., Maurer, T., 2017. Trends und Strategien in Logistik und Supply Chain Management: Chancen der digitalen Transformation. DVV Media Group GmbH, Bremen, 74 pp.

[7] Konovalenko, I., Ludwig, A., 2019. Event processing in supply chain management – The status quo and research outlook. Computers in Industry (105), 229–249.

[8] Luckham, D., Schulte, R., 2008. Event Processing Glossary: Version 1.1. Event Processing Technical Society. http://complexevents.com/wp-content/uploads/2008/08/epts-glossary-v11.pdf. Accessed 26 July 2021.

[9] Bensel, P., Fürstenberg, F., Vogeler, S. Supply-Chain-Event-Management : Entwicklung eines SCEM-Frameworks.

[10] Heusler, K.F., Stölzle, W., Bachmann, H., 2006. Supply Chain Event Management: Grundlagen, Funktionen und potenzielle Akteure. Wirtschaftswissenschaftliches Studium 35 (1), 19–24.

[11] Werner, H., 2017. Supply Chain Management: Grundlagen, Strategien, Instrumente und Controlling, 6. aktualis. u. überarb. Auflage ed. Springer Gabler, Wiesbaden, 548 pp.

[12] Onken, M., 2017. Internationale technische Standards, in: Müller-Mielitz, S., Lux, T. (Eds.), E-Health-Ökonomie. Springer Gabler, Wiesbaden, pp. 623–645.

[13] Tamm, G., Tribowski, C., 2010. RFID. Springer, Berlin [u.a.].

[14] 2017. EPCIS and CBV Implementation Guideline: Using EPCIS and CBV standards to gain visibility of business processes. https://www.gs1.org/docs/epc/EPCIS_Guideline.pdf. Accessed 26 July 2021.

[15] 2012. EPCIS – EPC Information Services: Prozess-Transparenz in Echtzeit, Köln. https://www.gs1-germany.de/fileadmin/gs1/basis_informationen/epcis_epc_informationservices.pdf. Accessed 26 July 2021.

[16] Stölzle, W., Schmidt, T., Kille, C., Schulze, F., Wildhaber, V., 2018. Digitalisierungswerkzeuge in der Logistik: Einsatzpotenziale, Reifegrad und Wertbeitrag: Impulse für Investitionsentscheidungen in die Digitalisierung - Erfolgsgeschichten und aktuelle Herausforderungen, 1. Auflage ed. Cuvillier, Göttingen, 147 pp.

[17] Bauer, D., Bauernhansl, T., Sauer, A., 2019. Enhanced Classification of Events for Manufacturing Companies in Supply Networks. Procedia CIRP (81), 87–92.

[18] Tröger, R., 2014. Supply Chain Event Management – Bedarf, Systemarchitektur und Nutzen aus Perspektive fokaler Unternehmen der Modeindustrie. Diss., Leipzig.

[19] Wiesner, O., Lauterbach, B., 2001. Supply Chain Event Management mit mySAP SCM (Supply Chain Management). HMD Praxis der Wirtschaftsinformatik (219), 65–71.

[20] Reiche, F., Hofstetter, J.S., Stölzle, W., 2009. Ereignisorientierte Steuerung von Lieferketten: Nutzen, aktueller Stand der Nutzung und Potenziale, 1. Auflage ed. Cuvillier, Göttingen, 77 pp.

[21] Levy, Y., Ellis, T.J., 2006. A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. Informing Science Journal 9, 181–211.

[22] Webster, J., Watson, R.T., 2002. Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly 26 (2), xiii–xxiii.

[23] Kovac, F., 2013. Untersuchung der Auswirkungen einer RFID-gestützten Bauzustandsdokumentation auf die Dokumentationsqualität in der Erprobungsphase: Am Beispiel ausgewählter Baureihen eines Automobilunternehmens. KIT Scientific Publishing, Karlsruhe.

[24] Böse, F., Piotrowski, J., Scholz-Reiter, B., 2009. Autonomously controlled storage management in vehicle logistics–applications of RFID and mobile computing systems. International Journal of RF Technologies 1 (1), 1–20.

[25] Werthmann, D., 2020. RFID-basierte Fahrzeugidentifikation und EPCIS-basierter Datenaustausch zur Verbesserung der Fahrzeugdistribution. Universität Bremen, Bremen.

[26] Schulte, C., 2013. Logistik: Wege zur Optimierung der Supply Chain, 6th ed. Vahlen, München, 750 pp.

[27] Souibgui, M., Atigui, F., Zammali, S., Cherfi, S., Yahia, S.B., 2019. Data quality in ETL process: A preliminary study. Procedia Computer Science 159, 676–687.

[28] Apel, D., Behme, W., Eberlein, R., Merighi, C., 2010. Datenqualität erfolgreich steuern: Praxislösungen für Business-Intelligence-Projekte, 2nd ed. Carl Hanser, München, Wien.

[29] Hinrichs, H., 2002. Datenqualitätsmanagement in Data Warehouse-Systemen, Oldenburg.

[30] Theodorou, V., Jovanovic, P., Abellò, A., Nakuçi, E., 2017. Data generator for evaluating ETL process quality. Information Systems 63, 1–21.

[31] Hamed, I., Ghozzi, F., 2015. A knowledge-based approach for quality-aware ETL process, in: 6th International Conference on Information Systems and Economic Intelligence (SIIE). 2015 6th International Conference on Information Systems and Economic Intelligence (SIIE), Hammamet, pp. 104–112.

[32] Helfert, M., Herrmann, C., Strauch, B., 2001. Datenqualitätsmanagement, St. Gallen.

[33] Rahm, E., Do, H.H., 2000. Data cleaning: Problems and current approaches. IEEE Data Engineering Bulletin 23 (4), 3–13.

[34] Bai, Y., Wang, F., Liu, P., 2006. Efficiently Filtering RFID Data Streams, in: CleanDB. CleanDB, Seoul, pp. 50–58.

[35] Helfert, M., 2000. Massnahmen und Konzepte zur Sicherstellung der Datenqualität, in: Jung, R., Winter, R. (Eds.), Data Warehousing Strategie. Erfahrungen, Methoden, Visionen. Springer, Berlin, Heidelberg, pp. 61–77.

[36] Blum, M.F., 2019. Der Digitale Schatten in der Auftragsabwicklung in der Einzel- und Kleinserienfertigung. Apprimus, Aachen.

Biography

Jokim Janßen (*1993) works as a project manager and scientific assistant at the Institute for Industrial Management (FIR) at the RWTH Aachen since 2019. Before that, he gained experience in supply chain management at Miele & Cie. KG and developed a strong background in event-based systems. In his current work as the group leader in supply chain management, he is managing projects in the context of digitalized supply chain management in research, consortia, and consulting projects.

Prof. Dr.-Ing. Dipl.-Wirt.-Ing. Günther Schuh (*1958) is the director of the Institute for Industrial Management (FIR) at RWTH Aachen University, holds the Chair for Production Systems at the WZL, and is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT in Aachen. Professor Schuh is a member of several supervisory boards and boards of directors.

Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

Procedure Model For Dimensioning And Investment Cost Calculation In An Early Factory Planning Phase

Leonard Rieke[1], Cihan Cevirgen[1], Alexander Gloy[1], Peter Nyhuis[1]
[1] *Institute of Production Systems and Logistics/ Leibniz University Hannover*

Abstract

Companies and their factories face constant change in today's world. Cost-intensive factory planning projects are being carried out in shorter intervals due to the increasing dynamism of the production environment. The related investments have a substantial impact on the liquidity of companies. Shorter production life cycles and changing consumer behaviour also require an adapted and more sustainable factory planning and cost estimation. However, especially in an early planning phase, available data and information are often uncertain and inaccurate. This effects in particular the outcome of the central dimensioning variables (operating resources, employees and area) for the planned factories. Incorrect dimensioning of these variables and thus of the associated costs can lead to substantial misinvestments. A holistic approach to obtain a reliable cost estimation of the factory project at an early stage is not yet available. This article therefore presents the development of a comprehensive procedure model for dimensioning and investment cost calculation in an early factory planning phase. For this purpose, relevant information and planning tasks with regard to dimensioning and cost estimation have to be identified first. Determined output values of the subsequent resource dimensioning represent the input values for the cost calculation. With the identification of surcharge factors, cost rates and calculation methods, the dimensioning variables, in particular the production area as the basis for the planned factory, can be estimated in terms of costs at an early stage.

Keywords

Investment cost calculation; procedure model; factory dimensioning; economic assessment; planning tool.

1. Introduction and problem definition

Factory planning can be described as a key factor for the economic success of companies [1]. Factory planning essentially involves planning the buildings, the production plant layout and the linking of the organizational units to each other including the material, personnel and information flow. Since factory planning lays the foundation for the entire production-side infrastructure, decisions in factory planning are often of a strategic and long-term nature. Accordingly, appropriate planning of different project contents is of essential importance for the manufacturing industry [2,3].

Due to their uniqueness and the complex positioning between cost, quality and time with often little or uncertain information, projects are generally subject to uncertainty [4]. Factory planning projects in particular represent a special challenge due to their long life cycle and significant investment costs. The early planning phase in factory planning projects plays a particularly important role. On the one hand, this is where the scope of action for each factory planning project is set [3,5]. On the other hand, the early planning phase presents planners with major challenges. Especially small and medium sized companies are lacking in sufficient resources, so that preliminary planning activities are lost in the operational business. In addition

DOI: https://doi.org/10.15488/12156

to that a reliable information database is missing [6]. However, reliable and accurate information are important basic conditions for factory planning projects [7]. In consideration of the fact that the cost influence is particularly high in the conceptual early planning phase and decreases with advancing planning progress, decision-supporting methods and approaches are especially relevant in an early phase [5,8]. Since the largest investments are determined at this point, it also means that the greatest influence on future project costs can be exerted at this early stage of planning [9]. This underlines the importance of economic efficiency assessments for decision support already at the beginning of the projects [5,9]. The dilemma of early factory planning quickly becomes apparent, since potential projects must be identified and selected with high uncertainties and inaccuracies without a secure planning basis and important economic assessments [10–12]. These uncertainties in the early planning phase significantly jeopardize the final economic planning of new factories [13]. If companies do not subject these projects to an appropriate assessment of economic efficiency, the risk of misinvestment increases, which in the worst case can lead to insolvency especially for small and medium-sized enterprises [14]. For this reason, it must be ensured that costs can be estimated at an early stage with sufficiently reliable data. Otherwise, there is a risk that calculations and reported costs are not sufficiently valid and thus damaging to the business [15]. In practice, these estimates are often based on individual experts and are thus subjective and insufficiently reproducible [6,16]. Various sources underline the omnipresent problem of finding suitable and practically applicable solutions to estimate costs despite high uncertainty in early planning phases [7,10,17,18].

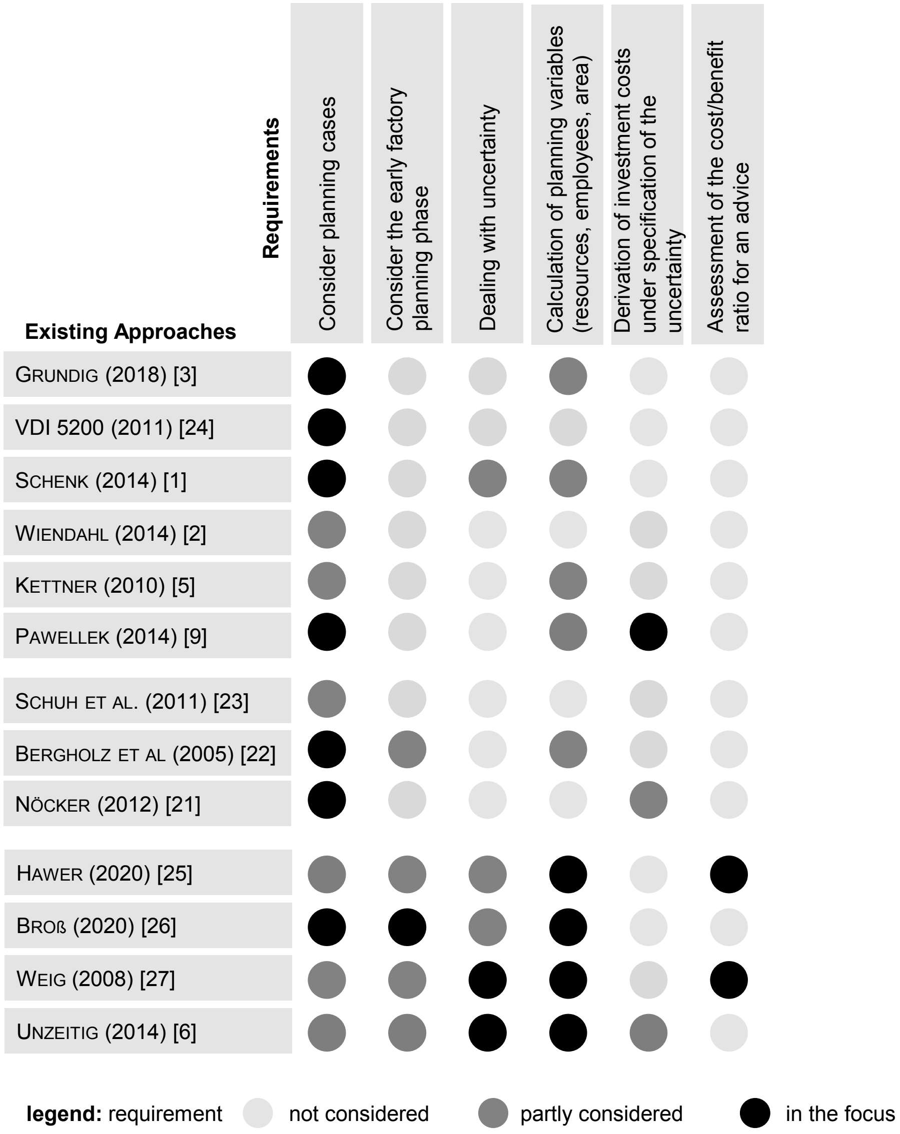
Initial preliminary work has shown that feasibility studies [19] and digital planning tools [20] can address this problem and provide added value. These initial approaches detailed the problem and outlined possible solutions, but did not yet introduce a structured approach to solving the problem. The hypothesis of this paper is therefore that a procedure model for reliable cost estimation is required in early factory planning in order to provide companies with effective economic decision support in the context of factory planning. To achieve that, this paper first summarizes the necessary requirements for solving the problem and evaluates existing factory planning procedures and approaches with regard to the fulfilment of these requirements. Based on the resulting research gap, a procedure model is presented that provides an approach to close this research gap based on the identification of necessary planning information, early dimensioning of the variables operating resources, employees and area by including surcharge-based assessment of investment costs. Subsequently, a digital planning tool developed on the foundation of this work is presented, which supports the user in a structured and reproducible procedure in the selection of the right planning project.

2. Derivation of requirements and literature review

The introduction and statement of the problem lay the basis for the identification of requirements for a procedure model for dimensioning and cost calculation in the early phase of factory planning. Certain requirements are identified as follows:

- Consider planning cases
- Consider the early factory planning phase
- Dealing with uncertainty
- Calculation of planning variables (operating resources, employees, area)
- Derivation of investment costs
- Assessment of cost/benefit ratio for an advice

With these basic requirements, existing approaches will first be considered in order to derive potential research need. For this purpose, both the classic factory planning approaches and specific approaches that focus on uncertainty and the early planning phase were selected (Figure 1). Numerous procedures for factory planning exist in the literature and practice, including [1,2,5,9]. Several approaches have been summarized in the VDI Guideline 5200, which divides the planning process into seven different planning phases. This guideline reflects the interdisciplinary character of factory planning [24]. The classic factory planning process consists of sequential planning phases that lead from the "rough to fine" to a gradually detailed planning status [2,3]. The VDI 5200 combines the design of logistical and technological processes with the building planning according to the fees for architects and engineers [2]. Overall, a distinction can be made between the planning cases greenfield, brownfield, demolition and revitalization [24]. In addition, there are approaches, such as condition-based factory planning, which enable a parallel and modular factory planning procedure. Here, the aim is to design the planning process in a way that is adaptable to each specific application case [21–23]. Apart from the classic approaches, there are other procedures in the field of factory planning that deal in particular with uncertainty and early planning phases [6,25–27].



Requirements **Existing Approaches**	Consider planning cases	Consider the early factory planning phase	Dealing with uncertainty	Calculation of planning variables (resources, employees, area)	Derivation of investment costs under specification of the uncertainty	Assessment of the cost/benefit ratio for an advice
GRUNDIG (2018) [3]	●	○	○	◐	○	○
VDI 5200 (2011) [24]	●	○	○	○	○	○
SCHENK (2014) [1]	●	○	◐	◐	○	○
WIENDAHL (2014) [2]	◐	○	○	○	○	○
KETTNER (2010) [5]	◐	○	○	◐	○	○
PAWELLEK (2014) [9]	●	○	○	◐	●	○
SCHUH ET AL. (2011) [23]	◐	○	○	○	○	○
BERGHOLZ ET AL (2005) [22]	●	◐	○	◐	○	○
NÖCKER (2012) [21]	●	○	○	○	◐	○
HAWER (2020) [25]	◐	◐	◐	●	○	●
BROß (2020) [26]	●	●	◐	●	○	○
WEIG (2008) [27]	◐	◐	●	●	○	●
UNZEITIG (2014) [6]	◐	◐	●	●	◐	○

legend: requirement ○ not considered ◐ partly considered ● in the focus

Figure 1: Assessment of factory planning approaches regarding identified requirements according to [19]

The classic literature often focuses on the procedure of factory planning itself as described in [1,3,24]. The factor uncertainty and the early planning phase are mentioned in the classic planning approaches, but neither GRUNDIG, KETTNER or SCHENK further elaborate on them, nor do they include them in their planning in advance. VDI 5200 and WIENDAHL only insufficiently name uncertainty associated with factory planning and accordingly do not address it further. Furthermore, the classic planning procedures only partially address the calculation of the basic dimensioning variables [1,3,9]. In PAWELLEK, however, the consideration of cost accounting can be emphasized [9]. The condition-based factory approach of SCHUH, BERGHOLZ or NÖCKER also only marginally considers the calculation of planning variables and the derivation of investment costs [21–23]. Uncertainty in the early factory planning phase is further considered by HAWER, BROß, UNZEITIG and WEIG. However, there are differences in the level of detail. HAWER uses a risk assessment analysis to identify the uncertainty factors [25]. BROß uses an approach with fuzzy logic and keeps this fuzziness also in his factors [26]. This also applies to WEIG, who partly uses expert estimates to make statements [27]. Ultimately, the approach of UNZEITIG can be highlighted, which successfully deals with uncertainty in detail, but in the end, similarly to the other approaches, does not provide a holistic approach to derive investment costs in an early planning phase [6].

Overall, the approaches only insufficiently link the requirements, especially the early planning phase, the dimensioning and the investment cost estimation derived from this. Specific approaches only consider single requirements and do not reflect upon the problem holistically. However, sub-points of individual approaches are considered useful and therefore will be included in the following elaborations.

3. Towards an approach for investment cost calculation in the early factory planning phase

On the basis of this background, a procedure model (Figure 2) was developed to dimension planned factories and to estimate their investment costs at an early stage. The developed model concentrates on the strategic planning cases greenfield and brownfield, since these two cases are most important for prospective adjustments of a factory. The entire procedure was developed by means of a three-stage research and analysis.

In the **first step**, planning tasks and planning information in the form of input and output parameters are considered in order to generally identify the required information for dimensioning and cost calculation in an early planning phase. Based on this, the main dimensioning variables (operating resources, employees, area) are described and appropriate calculation methods are derived in the **second step**. Due to the limited and uncertain data basis in the early planning phase, surcharge factors, cost factors and uncertainty factors are identified in the **third step** in order to be able to use them to perform an early cost estimate for the roughly dimensioned factory. The developed procedure model is developed in the manner that it can be used in an early planning phase at the very beginning of VDI 5200, where cost estimates are still insufficient or subject to a high degree of uncertainty (cf. Figure 2). By using process data as input parameters and including them in the calculation of the spatial view, the procedure model underlines the importance of synergetic factory planning to link production planning and building planning more closely. The individual steps taken to develop the resulting procedure model are explained in detail below.

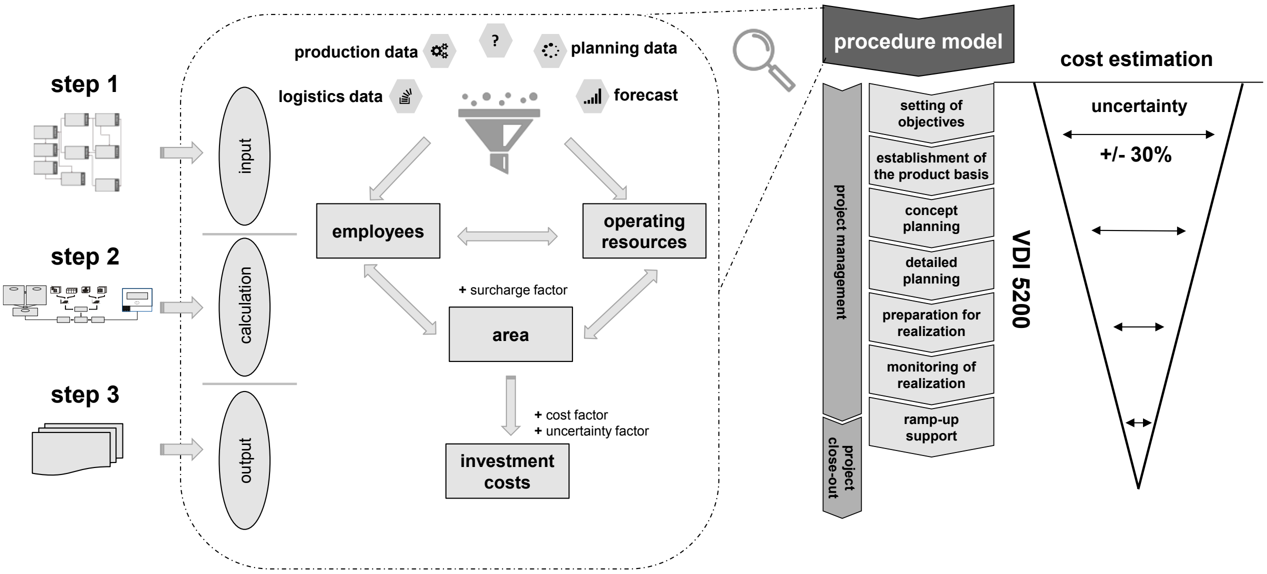


Figure 2: Procedure model for dimensioning and investment cost calculation in an early factory planning phase

3.1 Identification of planning information and planning tasks with regard to dimensioning

In order to dimension planned factories at an early stage, the required planning information and planning tasks first have to be defined. A common understanding has to be created concerning which information is available and at which point estimations may have to be made. For this purpose, a model (cf. Figure 3) was developed that describes in individual planning modules the procedure for dimensioning in an early planning phase.

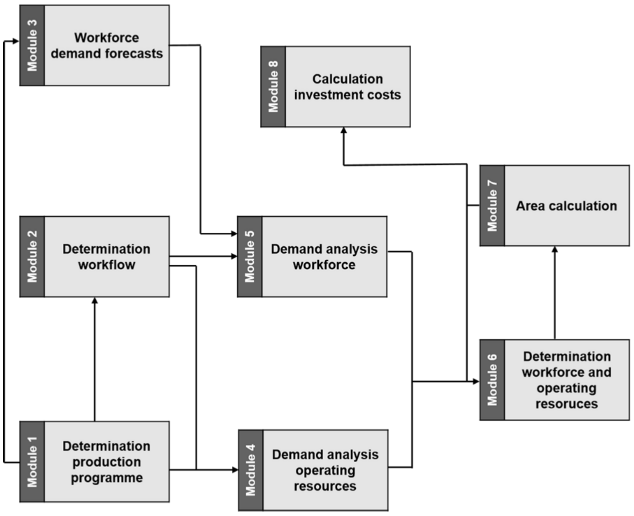


Figure 3: Step 1 - Identification of planning information and planning tasks in an early planning phase with regard to dimensioning

According to [23,25] the model consists of modules which have different input and output parameters within their planning tasks. In the early phase, for example, hardly any or no statements can be made about the material flow or detailed layout planning. In order to reduce structural complexity, this procedure was therefore limited to the basic dimensioning variables of operating resources, employees and area. Standard literature points out that these variables in particular should be taken into account in the early planning phase [2,3]. The first three modules "Determination production programme", "Determination workflow" as well as "Workforce demand forecast" serve to aggregate basic information regarding the three main dimensioning variables. Among other data, production programmes, sales forecasts, bills of materials, product variants, product characteristics, operating resource characteristics, work schedules and the organizational structure are analysed. The aggregated information from the first modules now serve as input parameters for the requirements analysis of operating resources and employees for the planned factory. With the assistance of a target/actual comparison of the available and required capacities, the requirements for employees and operating resources can thus be determined. These requirements in turn represent input information for the area calculation. Thus, area requirements for the planned factory are derived on the basis of the operating resources and employee requirements. The derived area is ultimately key input factors for the cost estimate in the course of the investment cost calculation. For this theoretically described procedure, calculation methods for dimensioning are now required that support the procedure and that are appropriate in the early planning phase.

3.2 Calculation methods for the key dimensioning variables of operating resources, employees and area

In the early planning phase of a factory, it is particularly important to consider the three main dimensioning variables of operating resources, employees and area. Due to this, these are briefly explained in this section, in order to subsequently derive and present adequate calculation methods (c.f. Figure 4).

The dimensioning variable **operating resources** is defined as technical work equipment for the fulfilment of a specific task in a work system. This technical work equipment is any "equipment (machines), plants, devices, measuring equipment, tools", which are used in a factory [3]. Their number and size with regard to the operating areas are of relevant importance for the development of a planning concept, especially with regard to the entire area programs. The type and number of operating resources required is largely defined by the respective product. For this purpose, demand figures (processing capacity) and availability figures (machine capacity) related to a specific period are compared with each other so that capacity deficits can be derived as a result, which form the basis for decisions on the dimensioning calculation of operating resources [3] (cf. Figure 4 according to [2]). Three possible outcomes can result from the comparison of the operating resources to the required demand of an already existing portfolio: A surplus, a shortage, and a balanced inventory [5]. From these correlations, measures for capacity adjustments can be derived, which have an effect on the dimensioning of the operating resources [2]. More detailed calculation methods for operating resources can be found in [3,5,28].

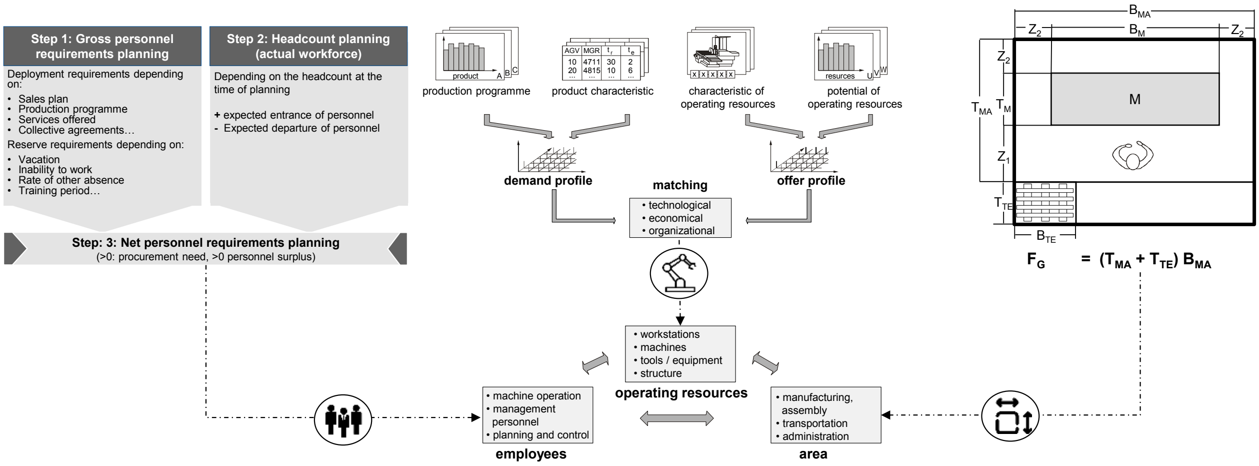


Figure 4: Step 2 - Calculation methods for the key dimensioning variables of operating resources, employees and area according to [2,29]

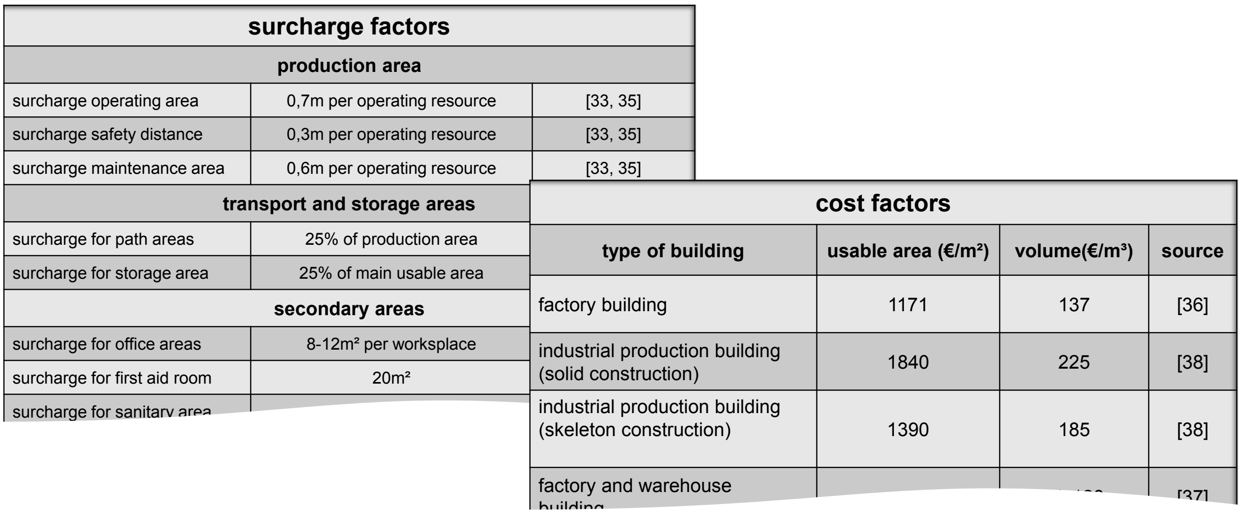
The dimensioning variable **employees** is determined by the number of workers required to perform all tasks in production and administration. The production programme is a useful reference point, as the corresponding demand is mapped and an estimate of the required number of workstations can be made [28]. In addition to the work areas, attention must also be paid to the installation of necessary functional rooms, such as toilets, showers and washrooms. The dimensioning variable employees can be derived from the required number of the workforce. A differentiated consideration of the number of personnel compared to the personnel requirements is necessary, which requires future changes in personnel planning. The planning of workforce requirements for a factory takes place in two ways, which should be carried out synchronously: On the one hand, qualitative personnel planning and, on the other hand, quantitative personnel planning [5]. The purpose of qualitative workforce planning is to match the required skills and knowledge of employees with the requirements of the work tasks [30]. It should be noted that companies are particularly restricted in short-term personnel planning by laws and collective agreement specifications. The required personnel demand has to be calculated for a future and long-term period and is mainly carried out for the areas of production, storage and administration in an early and rough planning phase [5]. Detailed calculation methods for employees can be found in [3,5,30] as well as in Figure 4 according to [29].

The **area** as the third dimensioning variable is mainly determined by the necessary area shares of the operating resources and the employees. Thus, the area becomes the central dimensioning variable, which also substantially determines the final layout of the building. According to VDI 3644, factory areas can be divided into main usable areas (especially production, storage and office areas) and secondary usable areas (in particular social areas and sanitary areas) as well as transport areas [31]. The determination of area requirements is considered to be an essential task of factory planning, as it attempts to adapt the design of areas to the necessary requirements for operating resources and employees. These adapted areas are realized in the layout and due to the usually limited available area supply, the available area must be constantly examined and reviewed for its active use with regard to the objective of maximum economic performance [3]. Detailed calculation methods for the area can be found in [2,3,5,32–34].

There are central interactions between the dimensioning variables, which makes a joint consideration essential. Thus, data and calculation results of the dimensioning of operating resources form the basis for individual area determinations, e.g. of the total machine working area. In particular, data regarding the dimensions and the capacity-related number of machines are crucial [2]. Figure 4 shows such a calculation-based dimensioning of the production area based on the actual machine footprint. In addition, the number of machines and the shift model determine the qualified employees required for the production process [2]. Furthermore, the degree of automation of the operating resources, for example, has an impact on the personnel requirements. In contrast, the introduction of a three-shift model ensures optimal utilization of the available capacities and equipment, but requires more personnel. Despite interactions between all dimensioning variables, changes in operation resources and personnel requirements always inevitably result in changes in area requirements. In the case of operating resources, it is the area reserved for the direct areas of the equipment and its periphery, and in the case of personnel for the indirect areas such as administrative areas or social areas. Since there is not enough information available at an early planning stage for the calculation methods outlined above, surcharge factors and cost as well as uncertainty factors in the dimensioning can provide useful assistance in the context of an investment cost estimate.

3.3 Identification and derivation of surcharge factors, cost factors and uncertainty factors

The results of the dimensioning are the basis for the final investment and cost decision [1]. The aim is to use a quantitative method to determine the profitable value of an investment from the given information and calculations. A summary of quantitative assessment methods for an investment can be found in [9]. However, due to the limited and uncertain data in an early planning phase, established and data-based cost calculation methods cannot be fully implemented. In order to still be able to provide a realistic picture of the planned factory in an early planning phase, surcharge factors, cost factors and uncertainty factors regarding the dimensioning of the area can be used. These are explained below in their practical application.



surcharge factors		
production area		
surcharge operating area	0,7m per operating resource	[33, 35]
surcharge safety distance	0,3m per operating resource	[33, 35]
surcharge maintenance area	0,6m per operating resource	[33, 35]
transport and storage areas		
surcharge for path areas	25% of production area	
surcharge for storage area	25% of main usable area	
secondary areas		
surcharge for office areas	8-12m² per worksplace	
surcharge for first aid room	20m²	
surcharge for sanitary area		

cost factors			
type of building	**usable area (€/m²)**	**volume(€/m³)**	**source**
factory building	1171	137	[36]
industrial production building (solid construction)	1840	225	[38]
industrial production building (skeleton construction)	1390	185	[38]
factory and warehouse			[37]

Figure 5: Step 3 - Identification and derivation of surcharge factors, cost factors, and uncertainty factors

First of all, valid area surcharge factors were identified depending on the operating resources and the employees in accordance with DIN 3644 for both the direct and the indirect areas. In the direct area, in addition to characteristic values regarding the machine dimensions, additional factors for individual area dimensions are important reference values, which allow a reasonable estimation of the total area. These supplements are, among others, values from the workplace guidelines (among others [35]) and are required in order to be able to carry out unproblematic operation, compliance with safety distances or necessary maintenance work. Commonly used methods in practice are the functional area calculation and the substitute area method in order to calculate the total production area on the basis of the machine base area with the help of surcharge factors (see also chapter 3.2). In addition, experience has shown that 25% of the production area can be allocated to path and transport areas [2]. In the indirect areas, the workplace guidelines and empirical values for surcharges in the area of offices and social areas are dependent on the number of employees. On the left, Figure 5 shows an excerpt of aggregated surcharge factors for calculating area space as a function of the dimensioning variables employees and operating resources.

In the final step of the procedure model, the area is included as a central input parameter in the investment cost estimate. Here, cost rates from various institutions that regularly analyse the cost structure of the construction industry are taken into account. The Federal Statistical Office, for example, presents the average costs at the time of approval of industrial buildings in its annually published reports on building permits [36]. Other sources [37,38] classify further cost rates and cost groups according to DIN 276 [39] on the basis of continuous random samples of new buildings. Depending on the required absolute area, the investment costs can now be estimated with these cost factors. It is possible to differentiate cost factors according to the construction method (skeleton construction, solid construction), construction quality (light, heavy) and depending on the required functionalities of direct and indirect areas. Cost factors for cleanrooms, for example, are many times higher than standard areas for industrial requirements. Figure 5 shows a section of possible cost factors on the right-hand side. These numbers represent general reference values from practice, which may differ with regard to the industry and company-specific requirements.

Cost estimates are increasingly subject to uncertainty due to the turbulent market environment. Since it is often not possible to calculate the exact requirements for the planned factory and, for example, the fluctuation of raw material prices or cost increases for technical building equipment validly in an early planning phase, uncertainty factors are often included in the investment cost calculation in practice. The data and information in this early planning phase are insufficiently accurate for exact calculations. Therefore, a variable uncertainty factor of 25-30% is added to the factory investment costs calculated in this procedure model. This factor represents a proven average value from practice [40–42].

3.4 Integration of the procedure model into a digital planning tool

The preliminary work and the developed process model support the dimensioning of factories in the planning phase at an early stage as well as its economic assessment. Surcharge, cost and uncertainty factors provide a solution in so far as information is often not yet completely available in an early planning phase. This procedure needs to be supported by a software tool in order to be able to carry out the dimensioning and estimation in a structured and reproducible way. The use of digital planning tools generally enables planning errors to be reduced and planning time to be cut while quality is increased [7,43]. Furthermore, the use of computer-aided tools can also contribute to mastering the complexity in the planning process [43]. Therefore, in parallel to the presented process model, a digital planning tool is developed that supports the user in a structured procedure and thus represents a reproducible decision support for the selection of the right planning project. In this context, the software-based cost calculator should also help to enable the early estimation of investment costs for planned factories for a wide range of users. Furthermore, by entering different input parameters, different scenarios for the future factory can be calculated and compared with each other. In this context, the main dimensioning variables are calculated according to the procedures and surcharges outlined above. On the basis of the calculated areas, investment costs for the individual factory areas can then be calculated.

4. Conclusion and Outlook

In a turbulent and uncertain market environment, factory planning with its interdisciplinary character becomes a complex and permanent task. In order to avoid misinvestments, planned factories should be subjected to a cost efficiency assessment at an early stage. However, in an early planning phase, information is often uncertain or not available. This paper therefore presents a procedure model that supports dimensioning and investment cost calculation in an early planning phase. In a three-stage procedure, planning information and tasks for the early dimensioning of factories are identified, calculation methods for the central dimensioning variables of operating resources, employees and area are derived and the resulting investment costs are estimated on the basis of surcharge, cost and uncertainty factors. To reduce complexity, a digital planning tool supports the user to estimate investment costs in a reproducible procedure and thus make future-proof decisions for the company. There is a need for further research in a more detailed consideration of uncertainty, e.g. with the aid of fuzzy logic, in order to take even closer consideration of the early planning phase. In addition, the interactions between the dimensioning variables have to be analysed in more detail in order to obtain a reliable planning basis for the subsequent cost calculation. Furthermore, certain target fields of factory planning, such as changeability or sustainability, could be integrated into the process model in order to be able to estimate the resulting investment costs for various scenarios.

References

[1] Schenk, M., Wirth, S., Müller, E., 2014. Fabrikplanung und Fabrikbetrieb: Methoden für die wandlungsfähige, vernetzte und ressourceneffiziente Fabrik, 2nd ed. Springer Vieweg, Berlin, Heidelberg.

[2] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2015. Handbook Factory Planning and Design. Springer Berlin Heidelberg, Berlin, Heidelberg.

[3] Grundig, C.-G., 2018. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 6., neu bearbeitete Auflage ed. Hanser, München.

[4] Khodakarami, V., Abdi, A., 2014. Project cost risk analysis: A Bayesian networks approach for modeling dependencies between cost items. International Journal of Project Management 32 (7), 1233–1245.

[5] Kettner, H., Schmidt, J., Greim, H.-R., 2010. Leitfaden der systematischen Fabrikplanung: Mit zahlreichen Checklisten. Hanser, München.

[6] Unzeitig, W., 2014. Methodik zur frühen Fabrikplanung bei Unsicherheiten. Dissertation, Graz.

[7] Bracht, U., Geckler, D., Wenzel, S., 2018. Digitale Fabrik: Methoden und Praxisbeispiele, 2nd ed. Springer Vieweg, Berlin.

[8] Frey, S.R., 1975. Plant layout: Planung, Optimierung und Einrichtung von Produktions-, Lager- und Verwaltungsstätten. Hanser, München.

[9] Pawellek, G., 2014. Ganzheitliche Fabrikplanung: Grundlagen, Vorgehensweise, EDV-Unterstützung, 2nd ed. Springer Vieweg, Berlin, Heidelberg.

[10] Rimpau, C., 2011. Wissensbasierte Risikobewertung in der Angebotskalkulation für hochgradig individualisierte Produkte. Herbert Utz Verlag, München.

[11] Herbst, P., 2014. Methode und Anwendung eines parametrischen Kostenmodells zur frühzeitigen Vorhersage der Produktentstehungskosten. Dissertation, Paderborn.

[12] Bundesministerium für Verkehr und digitale Infrastruktur (BMVI), 2015. Reformkommission Bau von Großprojekten: Komplexität beherrschen - kostengerecht, termintreu und effizient. Accessed 01.04.2022.

[13] Unzeitig, W., Rubesa, J., Schafler, M., Stocker, A., Ramsauer, C., Flasch, M., 2014. Eine Methode zur Berücksichtigung von Unsicherheit im Zuge der frühen Fabrikplanung in der Auftragsfertigung von Gesamtfahrzeugen. Elektrotechnik und Informationstechnik 131 (7), 212–218.

[14] Rautenstrauch, T., Müller, C., 2006. Investitionscontrolling in kleinen und mittleren Unternehmen (KMU). Z Control Manag 50 (2), 100–105.

[15] TGZ Bauökonomie, 2014. Datenquellen WU Hochbau: Deskription und Evaluation von Datenquellen für Wirtschaftlichkeitsuntersuchungen (WU) bei der Vorbereitung von Hochbaumaßnahmen des Bundes. Accessed 01.04.022.

[16] Souchoroukov, P., 2004. Improvement of cost estimating internal practice. Dissertation, Cranfield.

[17] Stewart, R.D., Wyskida, R.M., Johannes, J.D. (Eds.), 1995. Cost estimator's reference manual, 2nd ed. Wiley, New York.

[18] Niazi, A., Dai, J.S., Balabani, S., Seneviratne, L., 2006. Product Cost Estimation: Technique Classification and Methodology Review. Journal of Manufacturing Science and Engineering 128 (2), 563–575.

[19] Cevirgen, C., Rieke, L., Bischoff, L.-M., Nyhuis, P., 2021. Investment Feasibility Study for Factory Planning Projects. Journal of Production Systems and Logistics, Volume 1, Article 8.

[20] Rieke, L., Cevirgen, C., Nyhuis, P., 2021. Introduction of an Economic Assessment Approach for Factory Planning. Journal of Production Systems and Logistics, Volume 1, Article 9.

[21] Nöcker, J.C., 2012. Zustandsbasierte Fabrikplanung. Zugl.: Aachen, Techn. Hochsch., Diss., 2012, 1. Aufl. ed. Apprimus-Verl., Aachen, 287 pp.

[22] Bergholz, M., Schuh, G., 2005. Objektorientierte Fabrikplanung. Dissertation, Aachen.

[23] Schuh, G., Kampker, A., Wesch-Potente, C., 2011. Condition based factory planning. Prod. Eng. Res. Devel. 5 (1), 89–94.

[24] VDI - Verein Deutscher Ingenieure, 2011. VDI-Richtlinie 5200: Fabrikplanung- Planungsvorgehen / VDI-Guideline 5200: Factory planning - Planning procedures.

[25] Hawer, S., 2020. Planung veränderungsfähiger Fabrikstrukturen auf Basis unscharfer Daten. Dissertation, München.

[26] Broß, F.R., 2020. Dimensionierung indirekter Bereiche auf Basis unscharfer Daten. utzverlag, München.

[27] Weig, S., 2008. Konzept eines integrierten Risikomanagements für die Ablauf- und Strukturgestaltung in Fabrikplanungsprojekten. H. Utz, München.

[28] Burggräf, P. (Ed.), 2021. Fabrikplanung: Handbuch Produktion und Management 4, 2. Auflage 2021 ed. Springer Berlin, Berlin.

[29] Rationalisierungs-Kuratorium der Deutschen Wirtschaft, 1996. RKW-Handbuch Personalplanung, 3. Aufl. ed. Luchterhand, Neuwied.

[30] Drumm, H.J., op. 2008. Personalwirtschaft, 6., überarb. Aufl. ed. Springer, Berlin, Heidelberg.

[31] VDI - Verein Deutscher Ingenieure, 2010. VDI-3644 - Analyse und Planung von Betriebsflächen. Grundlagen, Anwendung und Beispiele.

[32] Podolsky, J.P., 1977. Flächenkennzahlen für die Fabrikplanung: Planungskatalog für metallverarbeitende Fertigungen. Beuth, Berlin, Köln.

[33] Kaufmann, H.-J., 1980. Flächenermittlung bei der Betriebs- und Werkstättenplanung: Fachbericht aus d. Institut für Fabrikanlagen, Universität Hannover. Beuth, Berlin.

[34] Rockstroh, W., 1982. Die technologische Betriebsprojektierung / Wolfgang Rockstroh ; Bd. 2: Projektierung von Fertigungswerkstätten, 2nd ed. Verlag Technik, Berlin.

[35] Bundesanstalt für Arbeitsschutz und Arbeitsmedizin - Ausschuss für Arbeitsstätten. ASR A1.2 Raumabmessungen und Bewegungsflächen.

[36] Statistisches Bundesamt (Destatis), 2019. Bauen und Wohnen - Baugenehmigungen / Baufertigstellungen Baukosten.

[37] BauGo, 2021. Baugebührenordnung; Preisindexzahl (Anlage 2 zur Baugebührenordnung) - Niedersachsen. Accessed 01.04.22.

[38] Kalusche, W., Herkel, S. (Eds.), 2021. Statistische Kostenkennwerte für Gebäude. BKI Baukosteninformationszentrum, Stuttgart, 932 pp.

[39] Deutsches Institut für Normung e.V., 2018. DIN 276 - Kosten im Bauwesen. Beuth-Verlag, Berlin.

[40] Siemon, K.D., 2012. Baukosten bei Neu- und Umbauten. Vieweg+Teubner Verlag, Wiesbaden.

[41] Werner, U., Pastor, W., Manteufel, T., Dölle, U., Frechen, F., Heinzerling, K., Wagner, K., 2020. Der Bauprozess: Prozessuale und materielle Probleme des zivilen Bauprozesses, 17., umfassend überarbeitete Auflage ed. Werner Verlag; Wolters Kluwer Deutschland GmbH, Köln, Hürth.

[42] Keldungs, K.-H., Baldringer, S. (Eds.), 2021. Architektenrecht: Praxishandbuch zu Honorar und Haftung, 7. Auflage ed. Werner Verlag; Wolters Kluwer Deutschland GmbH, Köln, Hürth.

[43] Bley, H., Fritz, J., Zenner, C., 2006. Die zwei Seiten der Digitalen Fabrik. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 101 (1-2), 19–23.

Biography

Leonard Rieke (*1995) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at Leibniz University Hannover since 2021. He previously studied industrial engineering (B.Sc., M.Sc.) at the Leibniz University Hannover (LUH).

Cihan Cevirgen (*1990) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at Leibniz University Hannover since 2018. He previously studied industrial engineering (B.Sc., M.Sc.) at the Leibniz University Hannover (LUH).

Alexander Gloy (*1998) has been studying industrial engineering (B.Sc., M.Sc.) at Leibniz University Hannover (LUH) since 2017. He has been a research assistant in the factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2021.

Peter Nyhuis (*1957) has been head of the Institute of Production Systems and Logistics (IFA) at Leibniz University Hannover since 2003. Prof. Dr.-Ing. habil. Peter Nyhuis was a member of the German Science Council (2014-2018) and has been Chairman of the Science Council of the AiF (GAG 3) since 2015. He is also a member of the German Academy of Technical Sciences acatech and the Scientific Society for Production Technology (WGP).

3rd Conference on Production Systems and Logistics

Analysis Of The Impact Of Lean Production Methods And Industry 4.0 Technologies On Sustainability And Flexibility

Fabian Dillinger, Alexander Kophal, Gunther Reinhart
Institute for Machine Tools and Industrial Management (iwb), Technical University of Munich, 85748 Garching b. Munich, Germany

Abstract

Today's manufacturing companies operate in a turbulent production environment characterized by globalization, mass personalization, and customer-specific product requirements. In this context, Lean Production and Industry 4.0 play an essential role for manufacturing companies. Both paradigms have farreaching production potentials for key performance indicators (KPI), such as time, cost, and quality. In addition to these KPIs, the Production's economic, ecological, and social sustainability and flexibility will also be important in the future. However, the influence of appropriate Lean Production methods and Industry 4.0 technologies on sustainability and flexibility has not yet been sufficiently researched. Therefore, this paper investigates the impact of Lean Production and Industry 4.0 elements on economic, ecological, and social sustainability and flexibility using a comprehensive literature review and an online survey with experts from science and industry. Thus, the results of this contribution support manufacturing companies to achieve their sustainability and flexibility goals with the help of Lean Production and Industry 4.0.

Keywords

Lean Production; Industry 4.0; Sustainability; Flexibility; Survey

1. Introduction

The industrial sector plays a crucial role in Europe. It contributes 75 % of European Union exports and 80 % of all innovations, making it a key driver of economic growth [1]. Nevertheless, with a 20 % share of global CO_2 emissions, the industrial sector is one of the main contributors to the worldwide effects of anthropogenic climate change [2]. Therefore, sustainability receives growing attention in production [3]. In addition, manufacturing companies face many complex influencing factors, such as volatile customer demands or short product life cycles [4,5], which require production flexibility. Since Lean Production and Industry 4.0 represent the two leading production paradigms of manufacturing companies [6], the question arises if Lean Production methods and Industry 4.0 technologies can meet the increasing demands for productions' flexibility and sustainability.

Lean Production is an established production philosophy that aims to reduce complexity in the value chain by eliminating all types of waste [7]. The characteristics of the concept are not limited to the reduction of waste and include the optimization of numerous production processes by implementing Lean Production methods [3]. Due to the advanced digitalization, further development of the production processes is required [8]. Regarding a fourth industrial revolution, Industry 4.0 brings significant changes to the economy, society, and environment. The goal is to enhance productivity by connecting all value chain participants to create a cyber-physical system using innovative technologies, such as predictive maintenance or artificial

DOI: https://doi.org/10.15488/12174

intelligence [9]. Both paradigms have far-reaching production potentials for key performance indicators (KPI), such as time, cost, and quality [10,11]. However, the impact of Lean Production methods and Industry 4.0 technologies on production's sustainability or flexibility needs to be further researched to support manufacturing companies in achieving their flexibility and sustainability goals by selecting and implementing the appropriate Lean Production methods and Industry 4.0 technologies. Therefore, this paper investigates the impact of Lean production methods and Industry 4.0 technologies on economic, ecological, social sustainability, and flexibility.

The following chapter sets the reference frame of the scientific fields and presents an introduction to Lean Production (2.1) and Industry 4.0 (2.1). Also, section 2.2 describes the relevant target dimensions of sustainability and flexibility. Chapter 3 analyzes the state of research (3.1) and identifies the research gap. To close this gap, a methodological approach is derived (3.2). This methodological approach forms the guideline for investigating the impact of Lean Production methods and Industry 4.0 technologies on sustainability and flexibility (chapter 4), which results will be discussed in chapter 5. The last chapter shows the limitations of the results and provides an outlook.

2. Fundamentals

2.1 Lean Production and Industry 4.0

After World War II, the Toyota Motor Corporation had to cope with low sales potentials on the Japanese automobile market. The lack of cost degression meant that mass production, according to Fordism, was not possible for Toyota [8]. Based on this initial situation, Taiichi Ohno designed the Toyota Production System (TPS), first described by Womak et al. [10] and is worldwide known as Lean Production. Lean Production aims to increase the production's economic efficiency by consistently and thoroughly eliminating all types of waste [6]. Moreover, the Lean Production methods aim to optimize production flow, realize a continuous value stream, and increase quality [12]. The two main principles of Lean Production are eliminating waste and continuous improvement, whereby employees should always be involved in the improvement process [13]. According to Dennis [14], Lean Production is based on four essential steps: The harmonization of the 4 M's (man, method, machine, and material), the optimization of the material flow, the introduction of the pull principle as well as the system improvement. To successfully implement the Lean Production approach with methods such as Kanban, value stream mapping, and Poka Yoke, the impacts on relevant target dimensions need to be known [15].

In addition to the Lean Production approach, Industry 4.0 was introduced in 2011 at the Hannover Messe in Germany [16]. Industry 4.0 is a technology-driven vision that aims to design smart factories and connect the physical and the cyber world with innovative technologies [17]. The so-called fourth industrial revolution is transforming the next generation of production systems by becoming intelligent, self-organized, decentralized, and flexible [18]. The digitization and networking of existing products, processes, and machines thus form the core of Industry 4.0 [19]. The goal is to organize the entire value chain, improve the efficiency of the production processes, and produce high-quality products and services. Further advantages are highly flexible mass production, reduction of complexity costs or coordination, and optimization of value chains in real-time [20]. Industry 4.0, therefore, seeks to realize the future factory by connecting employees and all physical resources of a production system, such as products, machines, transportation systems, and other objects, to achieve automated information exchange [9,21].

2.2 Target Dimensions: Sustainability and Flexibility

This study focuses on the impact of Lean Production methods and Industry 4.0 technologies on economic, ecological, and social sustainability and production's flexibility, which have become increasingly relevant in the industrial context. Target dimensions are needed to focus on long-term, strategic company goals rather than short-term improvements [23]. The term sustainability is used in various meanings [24]. The Brundtland Report presents the guiding principle of sustainable development [25]: "Sustainable development is a development that meets the needs of the present without compromising the ability of future generations to meet their own. (…) " [26]. The term can be specified by deriving three basic components of sustainability: An economical, ecological, and social dimension. These three dimensions form the triple bottom line and describe sustainable development as the simultaneous and equal implementation of economic, ecological, and social goals [27]. The dimensions can be characterized as follows [28]:

- Economic sustainability: Economic sustainability is the basis for the following dimensions and estimates the possibilities of a company to convert value creation potentials into competitive advantages and achieve long-term company continuity.
- Ecological sustainability: The ecological dimension includes the entrepreneurial influence on protecting and preserving the environment, and this requires a systematic reduction of ecological burdens and risks by companies.
- Social sustainability: The social dimension quantifies the social compatibility of entrepreneurial action and records the relationship construct with all stakeholders, such as employees and suppliers.

Nowadays, companies are confronted with volatile markets and globalization [9]. Therefore, a company's flexibility is increasingly becoming a strategic competitive advantage [29]. Flexibility is the ability of organizations to adapt to changing circumstances. The decisive factors are the timeframe and the extent to which companies react to changing situations, such as customer demands. The increasing complexity of the business environment is reflected in individualized demand and increased global competition. [29] Therefore, the adaptation of the production system is also necessary due to the modification of internal specifications and changes in external requirements [30].

3. State of the Art and Methodical Approach

3.1 State of the Art

A core principle of Lean Production is the elimination of waste, which also impacts sustainability by, for example, reducing costs, energy, and emissions [31]. Carvalho et al. [32] point out that not all waste elimination improves sustainability. The controversy is evident by investigating principles like Just in Time because operational costs are reduced through the effective use of warehouse space. At the same time, more frequent material handling leads to higher packaging material consumption and transportation emissions [32]. In contrast, little attention is paid to the relationship between Lean Production and social sustainability [31], although Lean Production methods, like Kaizen, impact employees' roles, require specific competencies [33], and increase the participation of its employees in decision-making [34]. A comprehensive study at the conceptual level was conducted by Varela et al. [35], who noted that the Lean Production approach is positively linked with sustainability and that, despite some barriers, synergies can be expected.

Also, according to the literature, Industry 4.0 makes it faster and easier to carry out economic decisions [36]. Digitalization influences ecological sustainability through the more efficient use of rare materials. Together with simplified disassembly, the waste of resources is counteracted and thus forms a basis for the circular economy [37]. However, social sustainability is affected in a conflicting way. Even though workers are acting in a safer environment, there is a risk that only highly skilled workers can handle and understand the new technologies, so that low-skilled workers may lose their jobs. [3]. Overall, Industry 4.0 benefits the economic [38–40] and ecological sustainability dimensions [39,41], but the impact on the social dimension remains questionable. The literature affirms that on a paradigm-level Lean Production and Industry 4.0 positively affect the flexibility in production, both individually and in combined applications [42,43]. This influence still needs to be explored on a detailed method and technology level.

According to the current state of the art, there is a positive correlation between Lean Production [35] respectively, Industry 4.0 [39], and the target dimensions of sustainability and flexibility [3]. The findings primarily relate to the overarching connections of the paradigms. However, individual methods and technologies are only presented as examples to visualize the results. Thus, there is a lack of in-depth research showing how individual methods and technologies influence the target dimensions. Also, according to Kabzhassarova et al. [3], there is a lack of empirical investigation of the literature-based findings. Therefore, it is essential to investigate the impacts of Lean Production and Industry 4.0 on sustainability and flexibility on the method and technology level.

3.2 Methodical Approach

In the following section, a systematic approach will be presented to examine the effects of Lean Production methods and Industry 4.0 technologies on economic, ecological, and social sustainability and flexibility in production. In the first phase, the Lean Production methods and Industry 4.0 technologies are collected and classified, resulting in an overview of the appropriate Lean Production methods and Industry 4.0 technologies. Afterward, an expert survey follows to derive the impact of both paradigms' elements on sustainability and flexibility (Figure 1).

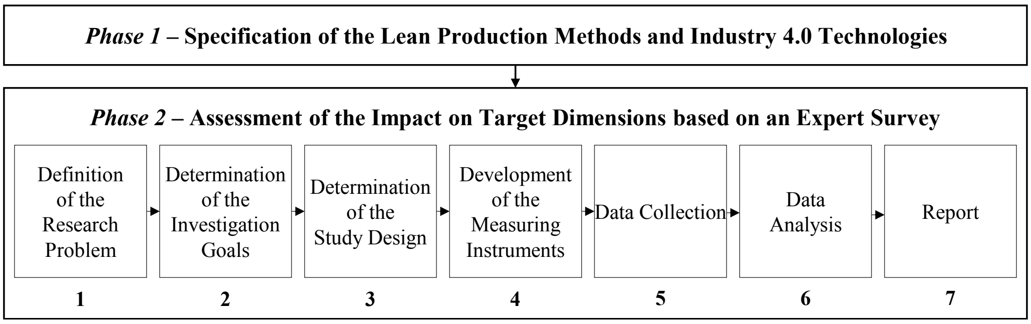


Figure 1: Methodical approach to attain the desired research aim (based on Kuß et al. [44])

The process for designing the research survey, as presented in Figure 1, can be divided into seven steps and is based on the work of Kuß et al. [44]. In the definition phase, the research problem is initially described as precisely as possible to specify the actual problem. Next, the study goals, which concretize and set the research task, are defined, and the study design is determined. The goals influence the study type, which must be considered to choose suitable methods and strategies. Once the structure is determined, measurement instruments must be developed to identify the characteristic attributes of the study subjects in the context. The data collection phase requires the most resources (time, human, financial). Here, possible errors, e.g., human weaknesses or technical problems, should be considered and the work status critically reflected. In the sixth step, statistical methods are used to analyze the collected data. Furthermore, the methods are essential for deriving conclusions that can be extrapolated from the results of a sample to the conditions in the corresponding population. The study's results are presented in the context of report writing or presentation of results, and the research questions should be answered.

4. Impact of Lean Production Methods and Industry 4.0 Technologies on Sustainability and Flexibility

4.1 Phase 1: Specification of the Lean Production Methods and Industry 4.0 Technologies

The selection of the Lean Production methods is based on Aull [45] and the VDI-2870 [47]. After the methods have been preselected by literature, an additional survey with participants from industry and science has been conducted to identify the relevant Lean Production methods [48]. Figure 2 provides an overview of twenty selected methods. In addition, the methods were classified according to Aull [45] into the categories logistics-oriented, employee-oriented, and quality-oriented [45].

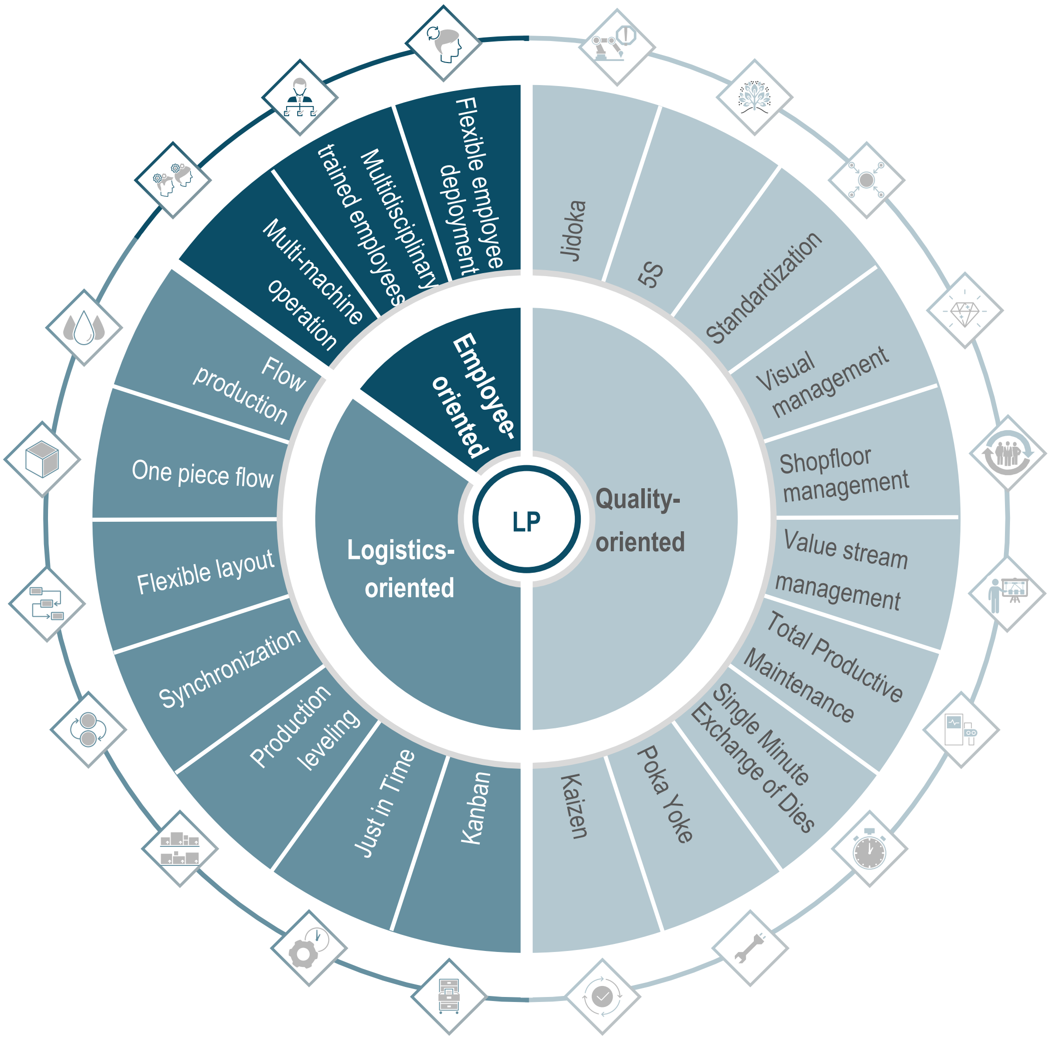


Figure 2: Collection of the Lean Production (LP) methods underlying this study [45]

According to Dillinger et al. [49], the Industry 4.0 technologies selection results from a comprehensive literature review, a use case analysis based on the Industry 4.0 platform of the Federal Ministry for Economics in Germany [50], and an expert survey. Based on the nine key technology of Rüßmann et al. [51], twenty-six Industry 4.0 technologies could be identified by Dillinger et al. [49]. The technologies were also separated into three main technology clusters, resulting from a mapping and clustering analysis using the software vosViewer [51]. The main clusters are smart data, smart operation, and smart interaction [53,52]. Figure 3 provides an overview of the twenty-six Industry 4.0 technologies considered in this study, visualizing the three clusters in the inner circle and the key technology fields in the middle circle. Finally, in this phase, the goals and descriptions for all the selected methods and technologies were formulated and summarized in a glossary given to the participants to ensure uniform understanding.

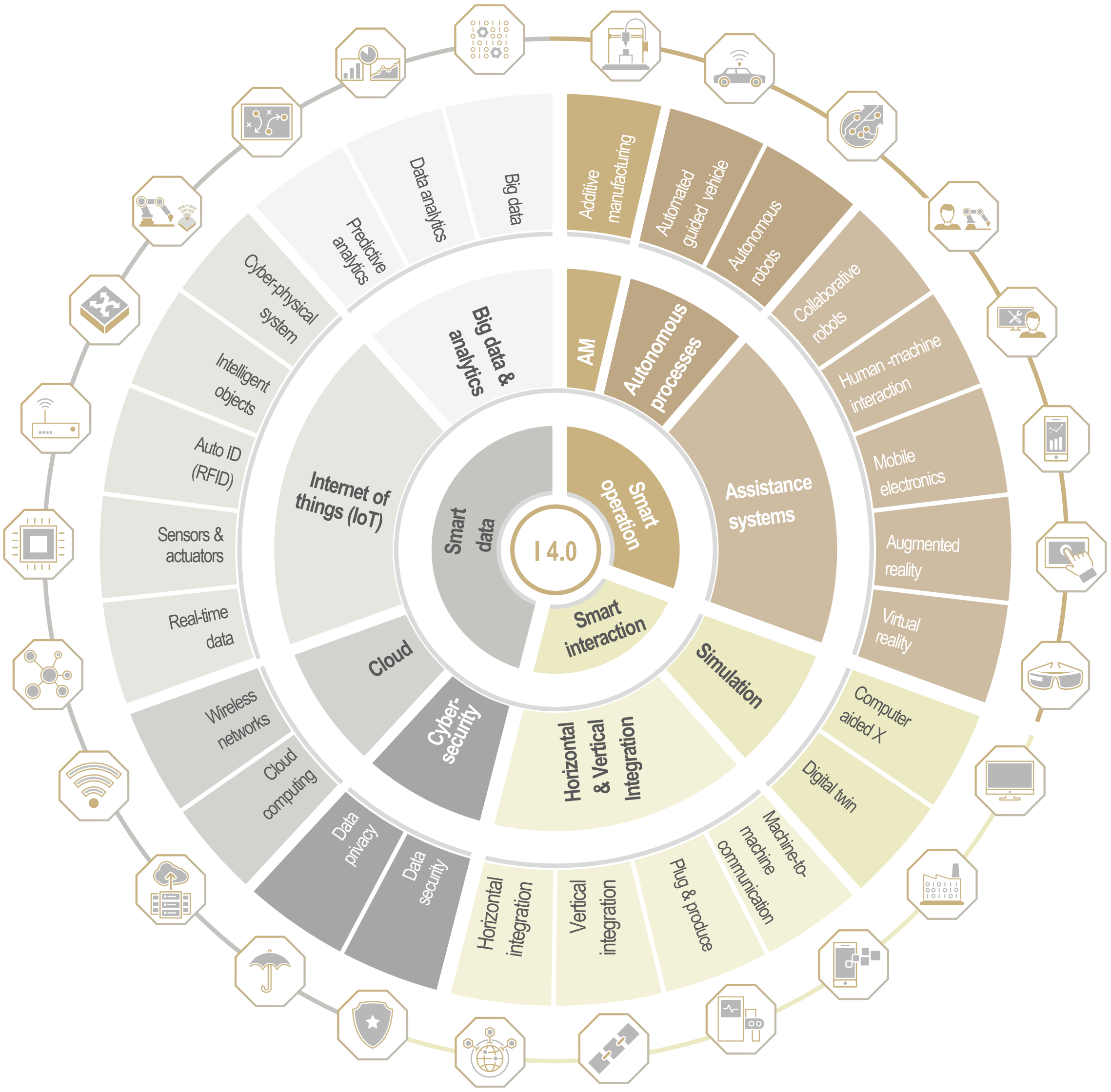


Figure 3: Categorization of the Industry 4.0 (I 4.0) technologies (according to Dillinger et al. [49])

4.2 Phase 2: Assessment of the Impact of Lean Production Methods and Industry 4.0 Technologies on Target Dimensions

In the second phase, an online survey with experts from science and industry was conducted to analyze the impact of Lean Production methods and Industry 4.0 technologies on economic, ecological, and social sustainability and flexibility in production.

The online study was designed according to established guidelines of empirical social research [55,54] and the systematic approach presented in Figure 1. It was conducted over three months and started in July 2021 with participating experts from production or production-related areas. In particular, people with knowledge of Lean Production and Industry 4.0 were required, such as production managers, production planners, or digital managers. In addition, management consultants and scientists were asked to strengthen the heterogeneity of the target group. With 32 experts, a representative cross-section of the German industrial landscape was reached. The study questions were answered using a seven-point Likert scale to determine the influence of Lean Production methods and Industry 4.0 technologies on sustainability and flexibility. The scale is sectioned from a very negative (-3) to a very positive impact (+3). In addition, participants had the option of choosing no effect (0) or could skip the question (k.A.), which ensures that the experts only assess the impact of methods and technologies that correspond to their expertise. The data analysis and the preparation of the report are summarized in Table 1.

Table 1: Assessment of the Impact of Lean Production methods and Industry 4.0 technologies

		Economic sustainability									Ecological sustainability									Social sustainability									Flexibility								
				Negative impact			No impact,	Positive impact					Negative impact			No impact	Positive impact					Negative impact			No impact	Positive impact					Negative impact			No impact	Positive impact		
		Ø	±	-3	-2	-1	0	1	2	3	Ø	±	-3	-2	-1	0	1	2	3	Ø	±	-3	-2	-1	0	1	2	3	Ø	±	-3	-2	-1	0	1	2	3
Lean Production Methods	Jidoka	**1.79**	0.98						●		**1.07**	0.93					●			**1.00**	0.95					●			**1.38**	1.03					●		
	5S	**1.27**	0.89					●			**1.30**	0.90					●			**1.61**	0.87						●		**1.06**	0.88					●		
	Standardization	**2.03**	0.80						●		**1.47**	0.85					●			**1.29**	0.96					●			**1.32**	1.42					●		
	Visual management	**1.33**	0.91					●			**1.00**	0.93					●			**1.50**	0.89						●		**1.81**	0.82						●	
	Shopfloor management	**1.57**	0.62						●		**0.97**	0.87					●			**1.58**	0.75						●		**1.45**	1.07					●		
	Value stream management	**1.87**	0.81						●		**1.43**	0.88					●			**1.13**	0.83					●			**1.27**	1.03					●		
	Total Productive Maintenance (TPM)	**1.93**	0.89						●		**1.50**	0.85						●		**1.16**	1.02					●			**1.35**	0.93					●		
	Single Minute Exchange of Die (SMED)	**1.93**	0.87						●		**1.03**	1.10					●			**0.53**	0.76					●			**1.83**	1.21						●	
	Poka Yoke	**1.73**	0.89						●		**1.17**	0.97					●			**1.23**	1.07					●			**0.94**	0.88					●		
	Kaizen	**2.03**	0.75						●		**1.33**	1.11					●			**1.55**	0.87						●		**1.35**	0.97					●		
	Kanban	**1.60**	0.92						●		**1.17**	0.91					●			**0.58**	0.75					●			**1.61**	0.83						●	
	Just in Time (JiT)	**1.87**	0.92						●		**0.80**	1.28					●			**0.23**	0.97				●				**1.23**	1.52					●		
	Production leveling	**1.56**	0.92						●		**1.10**	1.11					●			**0.93**	0.94					●			**1.18**	1.10					●		
	Synchronization	**1.67**	0.86						●		**1.10**	0.87					●			**0.82**	1.00					●			**1.54**	1.02						●	
	Flexible layout	**1.57**	0.76						●		**0.90**	0.88					●			**1.00**	1.05					●			**2.13**	0.98						●	
	One-piece flow	**1.47**	1.12					●			**1.07**	1.12					●			**0.48**	0.88				●				**1.63**	1.33						●	
	Flow production	**1.97**	1.02						●		**1.10**	0.94					●			**0.16**	1.11				●				**0.26**	1.41				●			
	Multi-machine operation	**1.80**	0.75						●		**0.53**	0.96					●			**0.26**	1.48				●				**1.31**	1.26					●		
	Multi-disciplinary trained employees	**1.50**	1.12						●		**0.97**	1.05					●			**1.87**	0.98						●		**2.45**	0.84						●	
	Flexible employee deployment	**1.73**	1.03						●		**0.79**	1.03					●			**1.48**	1.27					●			**2.53**	0.62							●
Industry 4.0 Technologies	Additive manufacturing	**1.45**	1.22					●			**1.28**	1.39					●			**0.27**	1.06				●				**2.20**	0.75						●	
	Automated guided vehicles	**1.73**	0.93						●		**1.00**	0.79					●			**0.42**	1.48				●				**1.65**	1.12						●	
	Autonomous robots	**1.70**	1.04						●		**0.85**	1.01					●			**0.42**	1.45				●				**1.63**	1.08						●	
	Collaborative robots	**1.70**	0.90						●		**0.59**	0.77					●			**0.90**	1.47					●			**1.73**	0.73						●	
	Human-machine interaction	**1.41**	0.95					●			**0.59**	0.67					●			**1.00**	1.48					●			**1.79**	0.76						●	
	Mobile electronics	**1.25**	0.74					●			**0.62**	0.72					●			**0.90**	1.35					●			**1.57**	0.90						●	
	Augmented reality	**1.07**	0.96					●			**0.86**	0.97					●			**1.23**	1.41					●			**1.48**	1.13					●		
	Virtual reality	**1.04**	0.98					●			**0.82**	0.97					●			**1.13**	1.41					●			**1.26**	1.19					●		
	Computer aided X (CAX)	**1.24**	0.97					●			**0.66**	0.92					●			**0.26**	1.05				●				**1.43**	1.02					●		
	Digital twin	**1.61**	1.08						●		**1.19**	0.94					●			**0.39**	1.07				●				**1.80**	1.01						●	
	Machine-to-machine communication	**1.32**	0.85					●			**0.93**	0.78					●			**0.42**	1.10				●				**1.73**	0.81						●	
	Plug & produce	**1.36**	0.85					●			**0.89**	0.82					●			**0.23**	1.09				●				**2.17**	0.83						●	
	Vertical integration	**0.88**	1.14					●			**0.70**	1.01					●			**0.31**	1.12				●				**1.36**	0.89					●		
	Horizontal integration	**1.28**	0.92					●			**0.79**	0.82					●			**0.27**	1.16				●				**1.40**	0.94					●		
	Data security	**0.48**	1.00				●				**0.07**	0.70				●				**0.97**	1.33					●			**0.03**	0.84				●			
	Data privacy	**0.41**	0.99				●				**0.04**	0.57				●				**1.20**	1.38					●			**-0.17**	0.73				●			
	Cloud computing	**1.29**	0.99					●			**0.34**	1.29				●				**0.28**	1.11				●				**1.43**	1.02					●		
	Wireless networks	**1.50**	0.98						●		**0.62**	0.93					●			**0.66**	1.06					●			**1.61**	0.90						●	
	Real-time data	**1.71**	0.84						●		**1.33**	1.11					●			**0.70**	1.16					●			**2.16**	0.72						●	
	Sensors & actuators	**1.26**	0.80					●			**0.67**	0.90					●			**0.33**	1.04				●				**1.36**	1.13					●		
	Auto ID (RFID)	**1.29**	0.80					●			**0.69**	0.79					●			**0.26**	1.01				●				**1.60**	0.99						●	
	Intelligent objects	**1.62**	0.79						●		**1.34**	0.84					●			**0.72**	1.28					●			**1.93**	1.01						●	
	Cyber-physical systems	**1.72**	0.87						●		**1.08**	1.03					●			**0.61**	1.26					●			**2.04**	0.82						●	
	Predictive analytics	**1.93**	0.84						●		**1.45**	0.85					●			**0.84**	1.17					●			**1.40**	1.05					●		
	Data analytics	**1.79**	1.09						●		**1.21**	1.11					●			**0.55**	1.04					●			**1.80**	0.95						●	
	Big data	**1.21**	1.19					●			**0.52**	1.23					●			**0.39**	0.94				●				**1.43**	1.26					●		

5. Results and Discussion

The survey results from Table 1 will be interpreted in this section, starting with Lean Production methods followed by the Industry 4.0 technologies. The participants attribute the highest positive impact on economic sustainability to the Lean Production methods. In this context, standardization (2.03) and Kaizen (2.03) show the highest positive scores. Concerning ecological sustainability, the impact of the Lean Production methods was weaker, and 95 % of the methods were rated with a low positive effect. The highest average was given to Total Productive Maintenance (TPM) (1.50), which indicates a positive impact followed by standardization (1.47). A reason why TPM has the highest positive impact on ecological targets is that it increases the overall equipment efficiency, which means that machine downtime can be avoided. When considering the impact on social sustainability, the ratings diverge between the methods. The methods 5S (1,61), visual management (1.50), shopfloor management (1,58), Kaizen (1.55), and multi-disciplinary trained employees (1.87) have a medium, positive impact. In contrast, the methods Just in Time (0.23), one-piece flow (0.48), flow production (0.16), and multi-machine operation (0.26) were rated as having no impact. With the latter four methods, the standard deviation must be considered. The standard deviation is higher compared to other methods. It describes a divergence because although the mean value suggests a neutral evaluation, both positive and negative effects were attested depending on the participants. In terms of flexibility, most methods received a low (55 %) to medium (35 %) positive rating. According to the participants, two outliers can be detected with flow production (0.26), which does not affect flexibility, and flexible employee deployment (2.53), which has a high, positive effect on flexibility in production.

The participants rate the importance of Industry 4.0 for economic sustainability by applying the technologies as predominantly low (54 %) to medium (38 %) positive. The highest rating is given to predictive analytics (1.93). Only the implementation of data security (0.48) and data privacy (0.41) is not considered to have any effect. Concerning the impact on ecological sustainability, the picture is uniform. Except for data security (0.07), data privacy (0.04), and cloud computing (0.34), which are not considered to have a significant impact, most of the technologies (88 %) are rated as having a low positive impact on ecological sustainability. When considering social sustainability, the participants rate the impact of Industry 4.0 technologies in part as having a low positive impact (54 %) and in part as having no impact (46 %). The first group primarily includes technologies that directly support employees, such as collaborative robots or human-machine interaction, whereas the second group includes digital twin or Auto ID. In terms of flexibility, most technologies have a medium positive impact (58 %). In particular, additive manufacturing (2.20), plug & produce (2.17), and real-time data (2.16) have the highest positive ratings. In contrast, the participants consider that the technologies data security (0.03) and data privacy (-0.17) have no or even a low negative impact on the flexibility in production.

When comparing the results of the twenty Lean Production methods and twenty-six Industry 4.0 technologies, it is noticeable that the participants assess the impact of Lean Production methods on sustainability more positively than the impact of the Industry 4.0 technologies. There is also a tendency toward a gradation from economic to ecological to social sustainability. According to the survey, Industry 4.0 technologies, in particular, positively influence production's flexibility. The literature review conducted by Kabzhassarova et al. [3] comes to a similar conclusion that Lean Production, in general, has the highest positive impact on economic sustainability and that the influences on the ecological and social dimensions cannot be determined. For Industry 4.0, they attest positive correlations for economic and ecological sustainability but cannot derive the effects on social sustainability. The expert survey shows that both approaches positively affect sustainability and flexibility by implementing their elements.

6. Conclusion and Outlook

This research paper provides a presentation of the impacts of Lean Production methods and Industry 4.0 technologies on economic, ecological, and social sustainability and flexibility in production. For this purpose, twenty relevant Lean Production methods and twenty-six Industry 4.0 technologies were identified, and an expert survey was conducted. The survey results show that the Lean Production methods and Industry 4.0 technologies have the highest positive impact on economic sustainability, followed by ecological and social sustainability. In the cross-paradigm comparison, it becomes clear that Lean Production methods' influence on sustainability is more positive than Industry 4.0 and its technologies. In particular, Industry 4.0 technologies positively impact the flexibility in production. Thus, the results of this contribution should be a first step to support manufacturing companies to achieve their sustainability and flexibility goals with the targeted selection of appropriate Lean Production methods and Industry 4.0 technologies. Future studies should deepen this research, and further experts should be consulted and use cases analyzed. Additionally, the impact of Lean Production methods and Industry 4.0 technologies on specific sustainability KPIs, e.g., CO_2 emissions or effects on employment contracts, should be investigated in detail.

References

[1] Blanchet, M., Rinn, T., Thaden, G.v., Thieulloy, G.d., 2014. Industry 4.0: The new industrial revolution. How Europe will succeed. Roland Berger Strategy Consultants GmbH, München.

[2] IEA, 2017. CO_2 Emissions from Fuel Combustion.

[3] Kabzhassarova, M., Kulzhanova, A., Dikhanbayeva, D., Guney, M., Turkyilmaz, A., 2021. Effect of Lean 4.0 on Sustainability Performance: A Review. Procedia CIRP 103, 73–78.

[4] Abele, E., Reinhart, G., 2011. Zukunft der Produktion: Herausforderungen, Forschungsfelder, Chancen. Carl Hanser Verlag, s.l., 262 pp.

[5] Dillinger, F., Kagerer, M., Reinhart, G., 2021. Concept for the development of a Lean 4.0 reference implementation strategy for manufacturing companies. Procedia CIRP 104, 330–335.

[6] Dillinger, F., Formann, F., Reinhart, G., 2020. Lean Production und Industrie 4.0 in der Produktion: Eine Studie zur Wechselwirkung und den gemeinsamen Potenzialen. ZWF 115 (10), 738–741.

[7] Monostori, L., Váncza, J., 2019. Towards living manufacturing systems. Procedia CIRP 93, 323–328.

[8] Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M., 2014. Industry 4.0. Bus Inf Syst Eng 6 (4), 239–242.

[9] Reinhart, G. (Ed.), 2017. Handbuch Industrie 4.0: Geschäftsmodelle, Prozesse, Technik. Hanser, München, 734 pp.

[10] Liebrecht, C., 2020. Entscheidungsunterstützung für den Industrie 4.0-Methodeneinsatz: Strukturierung, Bewertung und Ableitung von Implementierungsreihenfolgen. Dissertation. Shaker Verlag, Aachen, 1 Online-Ressource (169, LVI Seiten).

[11] VDI/VDE 4000 Blatt 1, 2021. Systematische Transformation und Evaluation von Produktionssystemen - Grundlagen.

[12] Liker, J.K., 2013. Der Toyota-Weg - 14 Managementprinzipien des weltweit erfolgreichsten Automobilkonzerns, 8th ed. Finanzbuch Verlag, München, 559 pp.

[13] Womack, J., Jones, D., Roos, D., 1990. The machine that changed the world: based on the Massachusetts Institute of Technology 5-milliondollar 5-year study on the future of the automobile. Rawson Associates, New York, 167 pp.

[14] Dennis, P., 2017. Lean Production Simplified, 3rd Edition, 3rd edition ed. Productivity Press; Safari, Erscheinungsort nicht ermittelbar, Boston, MA, 249 pp.

[15] Aull, F., 2013. Modell zur Ableitung effizienter Implementierungsstrategien für Lean-Production-Methoden. Zugl.: München, Techn. Univ., Diss., 2012. Utz, München, 232 pp.

[16] Perico, P., Mattilio, J., 2020. Empowering Process and Control in Lean 4.0 with Artificial Intelligence. Third International Conference on Artificial Intelligence for Industries (AI4I), 6–9.

[17] Tropschuh, B., Dillinger, F., Gärtner, Q., Korder, S., Bauer, H., Kagerer, M., 2021. Structure of a Socio-Technical Learning and Innovation Factory 269, 3–11.

[18] Gilchrist, A., 2016. Industry 4.0: The industrial internet of things. Apress, New York, NY, 250 pp.

[19] Bauernhansl, T., Hompel, M. ten, Vogel-Heuser, B., 2014. Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung, Technologien, Migration. Springer Fachmedien Wiesbaden, Wiesbaden, 648 pp.

[20] Bigliardi, B., Bottani, E., Casella, G., 2020. Enabling technologies, application areas and impact of industry 4.0: a bibliographic analysis. Procedia Manufacturing 42, 322–326.

[21] Tropschuh, B., Dillinger, F., Korder, S., Maier, M., Gärtner, Q., Vernim, S., 2021. Industrie 5.0 – ein menschzentrierter Ansatz. ZWF 116 (6), 387–392.

[22] Langlotz, P., Aurich, J.C., 2021. Causal and temporal relationships within the combination of Lean Production Systems and Industry 4.0. Procedia CIRP 96, 236–241.

[23] Gladen, W., 2014. Performance Measurement: Controlling mit Kennzahlen. Springer Gabler, Wiesbaden.

[24] Zimmermann, F.M., 2016. Was ist Nachhaltigkeit – eine Perspektivenfrage?, in: Zimmermann, F.M. (Ed.), Nachhaltigkeit wofür? Von Chancen und Herausforderungen für eine nachhaltige Zukunft. Springer Spektrum, Berlin, Heidelberg, 1–24.

[25] Holzbaur, U., 2020. Nachhaltige Entwicklung: Der Weg in eine lebenswerte Zukunft. Springer, Wiesbaden.

[26] Brundtland, G.H. (Ed.), 1987. Report of the World Commission on Environment and Development: Our Common Future.

[27] Steven, M., Klünder, T., 2018. Nachhaltigkeit schlanker Industrie 4.0-Netzwerke, in: Khare, A., Kessler, D., Wirsam, J. (Eds.), Marktorientiertes Produkt- und Produktionsmanagement in digitalen Umwelten. Springer Fachmedien, Wiesbaden, pp. 201–222.

[28] Hauff, M.v., 2014. Nachhaltige Entwicklung: Grundlagen und Umsetzung, 2nd ed. De Gruyter Oldenbourg, München.

[29] Zanker, C., Reisen, K., 2016. Stabilitäts- und Flexibilitätsanforderungen an Produktionssysteme, in: Kötter, W., Schwarz-Kocher, M., Zanker, C. (Eds.), Balanced GPS. Ganzheitliche Produktionssysteme mit stabil-flexiblen Standards und konsequenter Mitarbeiterorientierung. Springer Gabler, Wiesbaden, 13-37.

[30] Zäh, M.F., Bredow, M.v., Möller, N., Müssig, B., 2006. Bewertungsmethoden & Benchmarking - Methoden zur Bewertung von Flexibilität in der Produktion. Industrie-Management : Zeitschrift für industrielle Geschäftsprozesse 22 (4), 29–32.

[31] Tasdemir, C., Gazo, R., 2018. A systematic literature review for better understanding of lean driven sustainability. Sustainability 10 (7).
[32] Carvalho, A.C.V.d., Granja, A.D., Silva, V.G.d., 2017. A systematic literature review on integrative lean and sustainability synergies over a building's lifecycle. Sustainability 9 (7).
[33] Dillinger, F., Bernhard, O., Reinhart, G., 2022. Competence Requirements in Manufacturing Companies in the Context of Lean 4.0. Procedia CIRP 106, 58–63.
[34] Vinodh, S., Arvind, K.R., Somanaathan, M., 2011. Tools and techniques for enabling sustainability through lean initiatives. Clean Techn Environ Policy 13 (3), 469–479.
[35] Varela, L., Araújo, A., Ávila, P., Castro, H., Putnik, G., 2019. Evaluation of the Relation between Lean Manufacturing, Industry 4.0, and Sustainability. Sustainability 11 (5), 1439.
[36] Fathi, M., Nourmohammadi, A., Ghobakhloo, M., Yousefi, M., 2020. Production sustainability via supermarket location optimization in assembly lines. Sustainability 12 (11).
[37] Ghadimi, P., Wang, C., Lim, M.K., Heavey, C., 2019. Intelligent sustainable supplier selection using multi-agent technology: theory and application for Industry 4.0 supply chains. Computers & Industrial Engineering 127, 588–600.
[38] Erol, S., Jäger, A., Hold, P., Ott, K., Sihn, W., 2016. Tangible Industry 4.0: A Scenario-Based Approach to Learning for the Future of Production. Procedia CIRP 54, 13–18.
[39] Hofmann, E., Rüsch, M., 2017. Industry 4.0 and the current status as well as future prospects on logistics. Computers in Industry 89, 23–34.
[40] Stock, T., Seliger, G., 2016. Opportunities of Sustainable Manufacturing in Industry 4.0. Procedia CIRP 40, 536–541.
[41] Shrouf, F., Ordieres, J., Miragliotta, G., 2014. Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm, 697–701.
[42] Gallo, T., Cagnetti, C., Silvestri, C., Ruggieri, A., 2021. Industry 4.0 tools in lean production: A systematic literature review. Procedia Computer Science 180, 394–403.
[43] Mrugalska, B., Wyrwicka, M.K., 2017. Towards Lean Production in Industry 4.0. Procedia Engineering 182, 466–473.
[44] Kuß, A., Wildner, R., Kreis, H., 2018. Marktforschung: Datenerhebung und Datenanalyse, 6th ed. Springer Gabler, Wiesbaden, Heidelberg.
[45] Aull, F., 2012. Modell zur Ableitung effizienter Implementierungsstrategien für Lean-Production-Methoden. Dissertation, München.
[46] Busse, M., 2017. Implementierung Lean Management - Ein ganzheitliches Vorgehensmodell zur nachhaltigen Implementierung des Lean Managements in KMU. Dissertation, Cottbus.
[47] VDI 2870-2, 2013. Ganzheitliche Produktionssysteme. Methodenkatalog. Beuth Verlag GmbH, Berlin.
[48] Dillinger, F., Martl, N., Reinhart, G., 2021. Lean-Production-Methoden und Industrie-4.0-Technologien in der Produktion: Eine Studie zur Einführungsdauer und Relevanz. ZWF 116 (12).
[49] Dillinger, F., Messmer, C., Reinhart, G., 2021. Industrie-4.0-Technologiekreis für produzierende Unternehmen: Identifikation und Strukturierung relevanter Industrie-4.0-Elemente für die industrielle Produktion. Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (9), 639–643.
[50] Federal Ministry for Economic Affairs and Energy, 2021. Plattform Industrie 4.0. https://www.plattform-i40.de/PI40/Navigation/Karte/SiteGlobals/Forms/Formulare/EN/map-use-cases-formular.html. Accessed 28 November 2021.
[51] Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Engel, P., Harnisch, M., Justus, J., 2015. Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. BCG. https://www.bcg.com/de-de/publications/2015/engineered_products_project_business_industry_4_future_productivity_growth_manufacturing_industries. Accessed 14 July 2021.
[52] Lichtblau, K., Volker, S., Bertenrath, R., Blum, M., Bleider, M., Millack, A., Schmitt, K., Schmitz, E., Schröter, M., 2015. Industrie 4.0-Readiness. IMPULS-Stiftung, Aachen, Köln.
[53] Dillinger, F., Bernhard, O., Kagerer, M., Reinhart, G., 2022. Industry 4.0 Implementation Sequence for Manufacturing Companies. Prod. Eng. Res. Devel.
[54] Porst, R., 2014. Fragebogen: Ein Arbeitsbuch, 4., erweiterte Auflage ed. Springer VS, Wiesbaden, 210 pp.
[55] Häder, M., 2015. Empirische Sozialforschung. Springer Fachmedien Wiesbaden, Wiesbaden, 510 pp.

Biography

Fabian Dillinger, MBA & Eng., (*1991) is currently a PhD candidate at the Institute for Machine Tools and Industrial Management (iwb) at TUM. His research focuses mainly Lean 4.0 Management.

Alexander Kophal, B. Sc., (*1997) studied Mechanical Engineering at the Technical University of München. He is now a Master's student in the Mechanical Engineering program at the Technical University of Munich.

Prof. Dr.-Ing. **Gunther Reinhart** (*1956) chaired the Institute of Machine Tools and Industrial Management (iwb) at the Technical University of Munich.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2022

3rd Conference on Production Systems and Logistics

Throughput Analysis For Layout Optimisation Of Modular Conveyor Systems

Paul Aurich[1], Christian Böning[1], Malte Stonis[1], Ludger Overmeyer[1,2]
[1]*IPH – Institut für Integrierte Produktion Hannover gGmbH, Hollerithallee 6, Hannover, Germany*
[2]*Leibniz University Hannover, Institute for Transport and Automation Technology, An der Universität 2, Garbsen, Germany*

Abstract

In this paper, objective functions for the optimisation of modular conveyor systems will be introduced. Modular conveyor systems consist of conventional as well as modular conveyor hardware, which are arranged in form of matrix-like layouts. The aim of an ongoing research project is to provide small and medium-sized enterprises with a user-friendly decision support for the selection and planning of modular conveyor systems. For this purpose, the conveyor systems should be evaluated according to the objectives throughput and space requirement. Therefore, mathematical equations have been developed, which enable a fast and precise evaluation of layouts. The paper focuses mainly on the efficient calculation of the throughput. The result quality of the evaluation equations regarding the throughput was proven by a simulation of example systems.

Keywords

modular conveyor; conveyor system evaluation; throughput analysis; layout optimisation; logistics

1. Introduction

Conveyor systems are defined as technical devices that are used to automatically transport goods between two or more locations. Until now, conventional conveyor technology, such as belt, roller or chain conveyors, has been used in intralogistics systems such as warehouses, distribution centres or production. These conveyors are characterised by a high handling capacity. However, in case of complex material flows with many curves, junctions or intersections, they often require many additional routes, which are similar to the intersections on large motorways. This routing leads to an increased space requirement. At the same time, conventional conveyors are inflexible with regard to modifications, for example when adding new inputs and outputs or changing transport quantities.

In recent years, various modular conveyor systems were developed. Modular conveyor technology consists of functional components or modules that can quickly and flexibly be connected to each other via defined hardware and software interfaces. In this way, it is possible to adapt a conveyor system to changed requirements with minimal effort. In general, conveyor modules are characterised by a uniform aspect ratio. In addition, they can convey goods in multiple directions, in contrast to conventional conveyor technology, which can usually only convey in one dimension (forwards and backwards). In this way, different intralogistics functions can be realised, such as conveying, sequencing, buffering as well as infeed and outfeed. Another advantage is that the conveyor modules can be decentrally controlled. Examples of modular

DOI: https://doi.org/10.15488/12196

conveyor technology are the Celluveyor from cellumation [1] as well as the FlexConveyor and the GridSorter from flexlog [2]. In addition to these commercially available systems, there are also other modular conveyor systems such as the conveyor matrix from the research projects CogniLog and netkoPs (Figure 1), which are still being developed [3,4].

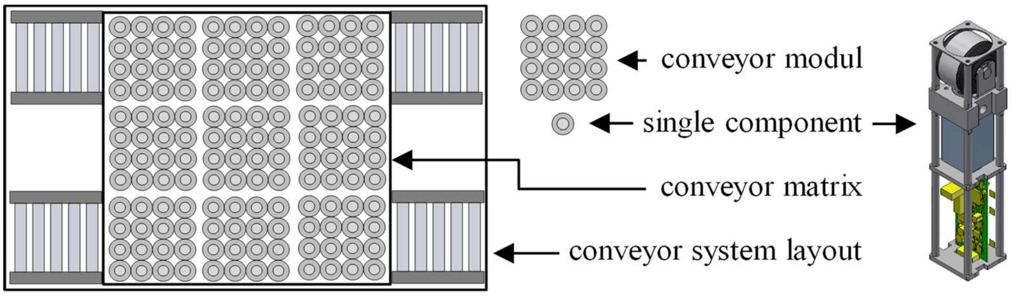


Figure 1: components of a modular conveyor system layout

Due to the two-dimensional transport, modular conveyor technology makes it possible to realise complex material flows in a very confined space. Due to the novelty of the systems, however, companies and logistics planners do not have substantial experience in planning and setting them up. For example, the question arises how many conveyor modules are necessary for the transport of a defined quantity of goods and how these must be positioned considering given inputs and outputs. To support this planning problem, a method for layout optimisation of modular conveyor systems is developed within the research project OptiLay – "Automated creation of optimised conveyor system layouts for modular conveyor systems" [6,5]. Optimisation algorithms are used for the placement of the conveyor modules. To evaluate the computer-generated layouts, a quantitative calculation or measurement of objective criteria is necessary. The evaluation must also be as fast and automated as possible in order not to slow down the optimisation. This paper presents an approach for the evaluation of modular conveyor systems for the objective criteria throughput and space requirement.

2. Related Research

The calculation and evaluation of the objective criteria space requirement can be performed with simple analytical formulas. Preliminary work on this is provided by SHCHEKUTIN [7]. The throughput of a conveyor system with only one material flow can also be calculated easily. If there are several material flows between different sources and sinks, complex intersection situations arise depending on the layout, for those no simple analytical approach exists to determine the total throughput. DALLERY AND GERSHWIN [8] as well as LI [9] provide exact solution approaches for simple intralogistics conveyor systems and machines with a steady-state distribution of the transport goods. These approaches are not suitable for complex conveyor systems with many conveyor modules, as the individual intersection situations have to be modelled in a complex process [10]. There are various approximation methods for the throughput calculation, which are based on the decomposition of the system into subsystems. For each subsystem, a throughput is calculated. The subsystems are then reconnected and the throughput of the entire system is estimated. ARNOLD provides a basic framework for the calculation of various subsystems [11]. SCHMIDT and JACKMANN developed a decomposition approach for recirculating conveyor systems with blocking before service. GAO ET AL. introduced a decomposition approach for multiple material flows [10]. They developed an algorithm for the decomposition into subsystems. The throughput analysis is based on their previous research [12]. The approach was applied to three conveyor systems in which only simple intersections and merges occur. The analysed systems were simulated for comparison, and a calculation correctness of 90 % could be confirmed. SHECHKUTIN also developed a layout optimisation approach for modular conveyor systems [7,13]. The study is mainly focused on the results of the research project netkoPs. The throughput calculation is based on the detached calculation of the throughput time per material flow. Here, the transport times for the used conveyors are accumulated. If the material flow runs through an intersection point or a modular conveyor matrix, the transport time of this transport route is offset by a factor. The factor is calculated for each intersection situation based on the size of the modular conveyor matrix. It is assumed that a larger conveyor matrix leads to a simplification of the conflicts. The conflicts themselves are not calculated concretely.

In addition to analytical modelling, simulation is used to evaluate the throughput of conveyor systems. Extensive simulation software is available for this purpose, such as AnyLogic or Plant Simulation. There are many throughput analyses in the literature that have been carried out using simulation. With the help of simulation, almost any complex conveyor system can be analysed. However, complex and time-consuming modelling is necessary, which has so far been done manually for the most part, since complex control procedures have to be implemented (routing, blocking prevention). Due to the modelling effort, simulation is not suitable for a quick calculation of the throughput in a layout optimisation. MAYER developed a routing method for the FlexConveyor [14]. Within the study, a throughput analysis was carried out by simulation for different layouts of the FlexConveyor (e.g. straight conveyor, line sorter, circles and circles with intersections). The aim was to check the routing approach. SEIBOLD developed and simulation-based validated a routing method for the GridSorter, focusing on avoiding deadlocks [15]. The GridSorter consists mainly of the components of the FlexConveyor, but is characterised by a uniform transport direction. With the help of the GridSorter, goods can be sorted between different lanes. KRÜHN [16] and SOHRT [17] also developed routing methods but for the CogniLog respectively netkoPs conveyor matrix. The routing methods were also validated with simulation. KRÜHN used a reservation logic to avoid blockages or deadlocks. SOHRT, on the other hand, developed a time-window-based approach.

In summary, none of the existing approaches meets the previously described requirements for direct application in the context of an optimisation (computing time and automated modelling). Therefore, the approach presented below was developed. Of course, existing approaches such as decomposition were adopted in the design process.

3. Representation of the optimisation problem

The optimisation problem is formulated as an extended quadratic assignment problem. Accordingly, the area in which the conveyor modules will be arranged is covered with a grid of uniform squares. This results in a discrete coordinate system of cells. The conveyor modules also have a square shape. The cells of the base area and the conveyor modules must be of equal size. Accordingly, only conveyor modules with the same dimensions can be combined within a layout. The sources and sinks of the conveyor system are adjacent to the footprint. Each material flow is defined by a type of specific goods, a transport quantity, a source and a sink. Each type of good has dimensional attributes in form of a horizontal and a vertical length. Whereby the horizontal length always reflects the longer side length of a good. This means that if a good is transported in a horizontal direction (e.g. from east to west in plan view), the longer edge is parallel to the direction of flow. Accordingly, the shorter edge is parallel to the flow direction when a good is transported in a vertical direction. If several material flows run parallel respectively together in one path section, they are combined with regard to their attributes. The transport quantity is summed up. The dimensions of the goods are converted into quantity-weighted average lengths.

Based on the arrangement of the conveyor modules, a graph of the conveyor system is also generated. This can be used to calculate the transport routes. For this purpose, the conveyor modules have a rotation attribute in addition to a position attribute. In this way, it can be checked whether the inputs and outputs of the conveyor modules are adjacent to each other and thus if a transport is possible. If this is the case, the nodes of the conveyor modules are connected by an edge. The transport path of a material flow can be determined by common path-finding algorithms. Figure 2 a) shows an exemplary conveyor system layout. Figure 2 b) illustrates the graph derived from it. Here it is important to consider that a conventional conveyor is represented by three nodes and a modular conveyor by up to five nodes. Each conveyor has a node for its centre point and nodes for the transitions to other conveyors. The adjacent transitions are aggregated into one node.

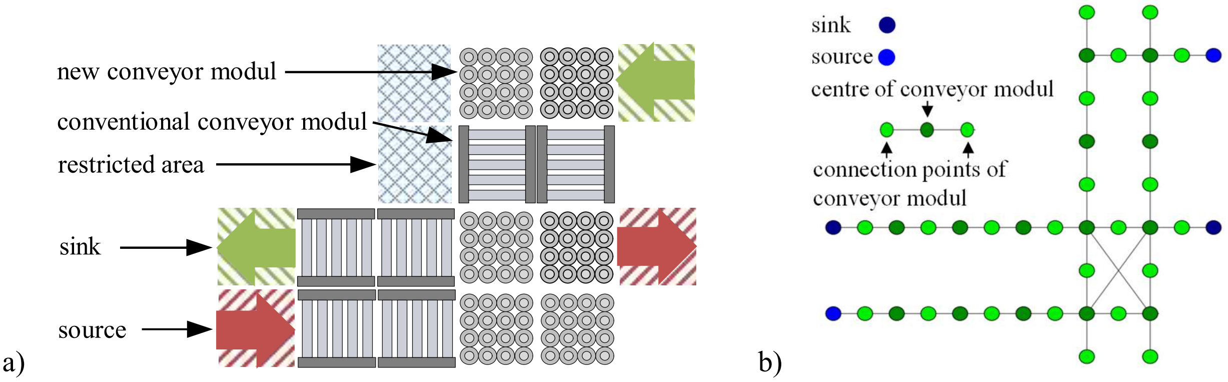


Figure 2: a) conveyor system layout, b) graph representation

4. Evaluation of conveyor systems

In the following section, the equations for the evaluation of conveyor systems are presented. In addition to calculating the objective values for each objective Z, it is also important to normalise them. This is necessary because the values of the objectives do not have the same scaling. For example, the summation of a throughput with 1,000 pieces per hour and a space requirement of 20 square metres would mean, that a change of the required space would have almost no influence on the sum of the objective values. There are several approaches to the normalisation of objective values. HARMONOSKY and TOTHERO developed a method in which each sub-value of an objective (e.g. the throughput of a single material flow) is divided by the sum of all sub-values for that objective [18]. SINGH and SINGH developed a procedure in which the normalisation is performed with the help of a multi-stage calculation process that includes standard deviations and mean values of sub-values [19].

Within this research project, the evaluation formulas of the objective criteria were designed in such a way that a percentage value between 0 and 100 is returned. Thus, the individual objective values neither have to be scaled nor normalised in order to enable a comparison. In most cases, the percentage value is computed by calculating the ratio of the best-known value to the current value of an objective. The disadvantage of the method is that when a new best objective value is found, all previously calculated values of the same objective must be updated. If these values are not directly needed to control the optimisation method, the update can also be done at the end of the optimisation. The weighting of the objectives is possible without restrictions.

4.1 Throughput

In order to calculate the throughput λ^k of a conveyor system k, the decomposition approach is also used. For this, the transport paths of the material flows are first checked for intersections in the graph. A conveyor module f is an intersection u if two or more material flows do not use the same adjacent conveyor modules. Accordingly, a parallel transport on a straight line or in a curve is not considered as an intersection unless the transport direction is opposite. This definition results in the intersections shown in Figure 3. Case B represents only one possible instance of an intersection over several conveyor modules, which is described below. There are no other intersection cases beyond the ones shown. A deviating number of material flows is mapped via a virtualisation of material flows. This means that material flows that pass through an intersection completely in parallel are combined into a virtual material flow based on their attributes.

In the throughput calculation, only the intersections are considered afterwards. In general, the bottleneck in a linked system determines the throughput of the system. In the case of the conveyor system, an intersection with the lowest throughput determines the throughput of the entire conveyor system (equation 2). This only applies as long as all material flows within a conveyor system intersect each other. If there are several independent material flows or material flow systems in a conveyor system, the bottleneck must be determined for each subsystem. In this case, the total throughput is the sum of the bottleneck throughputs of the subsystems.

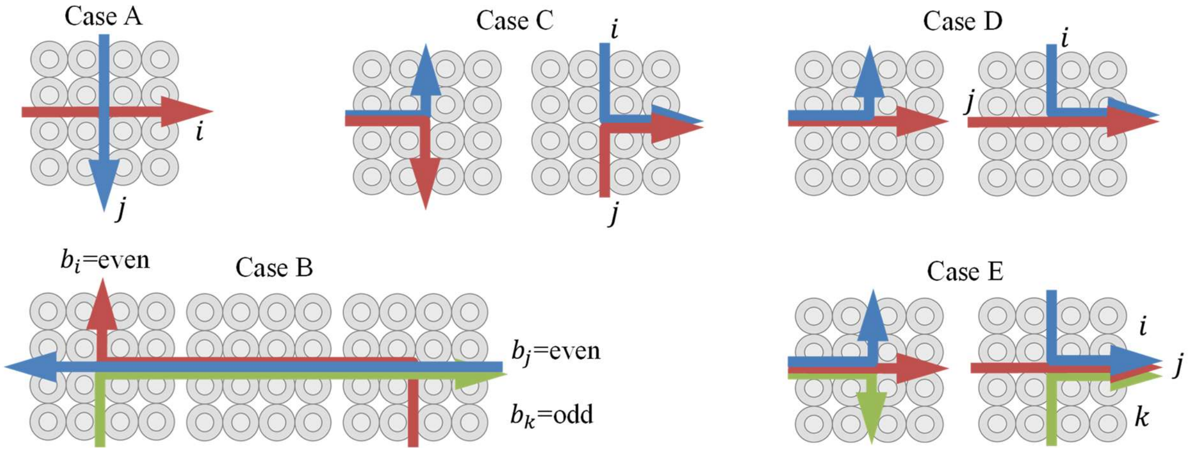


Figure 3: types of intersection scenarios

f: conveyor module, with $f \in F$

u: Conveyor module on which material flows intersect, with $u \in U$ und $U \subseteq F$

λ_u: throughput of an intersection u

λ^q: throughput of a conveyor system q

λ^{best}: best-known throughput

$$\lambda_u = \frac{t_u}{3600} \tag{1}$$

$$\lambda^q = \min(\lambda_u)\ \forall\ u \tag{2}$$

$$Z_\lambda = \frac{\lambda^q}{\lambda^{best}} * 100 \tag{3}$$

In the following subsections, evaluation formulas for all relevant intersection situations are presented. Basically, these are based on the calculation of the cycle time t_u. For this purpose, the quantities of goods a_i of the material flows i are first set in relation to each other in order to calculate a batch size n_i. For example, two material flows with 400 and 600 pieces per hour correspond to batches of 2 and 3 goods per load cycle. A load cycle represents the sequence of transport movements on a conveyor that are necessary to process the batches. The number of load cycles results from the greatest common divisor of the quantities of the material flows. Dividing the quantity of a material flow by the number of load cycles gives the batch size. Then the transport time t_u of each load cycle of the intersection is calculated. This calculation is based on the accumulation of transport distances. Distances are, for example, the conveyor module length s_u or the respective length of the goods to be transported l_{iH} and l_{iV}. The sum of the distances is then divided by the conveying speed v_u. The following parameters are used for the calculations:

$i; j; k$: material flow with specific good ($i \neq j \neq k$)

a_i: target quantity of goods [parts/hour]

n_i: batch size of material flow i

l_{iH}: dimension of a good of material flow i in horizontal direction (H)

l_{iV}: dimension of a good of material flow i in vertical direction (V)

i': virtual material flow resulting from combination of several material flows with same transport direction

$\overline{l_{i'H}}$: quantity-weighted dimension of a virtual good from the combination of several material flows i in horizontal direction (H)

$\overline{l_{i'V}}$: quantity-weighted dimension of a virtual good from the combination of several material flows i in vertical direction (V)

$\overline{n_{i'}}$: batch size of the virtual material flow

v_u: transport velocity of a conveyor module u

s_u: length of a conveyor module u

t_u: time for a load cycle of all material flows on a conveyor module $\sum t_{ui}$

t_{ui}: time in which a batch n_i of material flow i is transported on a conveyor module u

4.1.1 Basic intersection of two material flows (Case A)

The most basic form of intersection occurs when two material flows pass a conveyor module where the transport direction is offset by 90°. See case A Figure 3. The time of the load cycle is composed as follows. First all goods of the lot of material flow i pass the intersection. For this purpose, the corresponding lengths are added up. The lot size is multiplied by the length of the goods, and the distance across the intersection is added. This results in the total length that must be moved so that all goods of the batch pass the conveyor. This is then also done for the second material flow j, which flows in a 90° rotated direction.

$$t_u = \frac{(s_u + n_i * l_{iH}) + (s_u + n_i * l_{jV})}{v_u} \tag{4}$$

4.1.2 Intersection over several conveyor modules (Case B)

The extension of the basic intersection is already the most difficult intersection situation in terms of controlling a conveyor system. It is an intersection of material flows with opposite directions. To avoid colliding in the intersection, the goods must be stopped before entering. Like in other intersections, the goods of the material flows must wait for each other. Accordingly, all conveyor modules are blocked in the intersection situation. In order to calculate the load cycle, it must also be determined whether the goods change direction in the intersection, as this has an influence on the travel distance. If goods change direction with an even number, then they flow out of the crossing area in the same direction as they came in (equation 5a). If the number of direction changes is odd, the horizontal and vertical length of the goods must be considered once when calculating the transport distance (formula 5b).

h: number of conveyor modules f in an intersection situation in opposite directions

b_i: number of direction changes of a material flow in the intersection situation

$$t_{ui} = \frac{(h * s_u + n_i * l_{iV})}{v_u} \forall\, b_i \text{ or } t_{uj} = \frac{(h * s_u + n_j * l_{jH})}{v_u} \forall\; b_j\text{, with } b_i, b_j = \text{even number} \tag{5a}$$

$$t_{uk} = \frac{\left(h * s_u + n_k * \left(\frac{l_{kH}}{2} + \frac{l_{kV}}{2}\right)\right)}{v_u} \forall\, b_k = \text{odd number} \tag{5b}$$

$$t_u = \frac{\sum t_{ui}}{v_u} \tag{6}$$

4.1.3 Merging of two opposing material flows (Case C)

Previously, intersections were described where the material flows are independent of each other before and after passing the intersection. In addition, there are intersections where material flows are merged or separated. In the first of these cases, two material flows moving in opposite directions are merged and both make a turning movement. In the second case, two material flows come from the same direction and are separated by a turn. The equation for the separation is identical to the one for the merging.

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2} + \frac{s_u}{2} + \frac{l_{iH}}{2}\right) + n_j * \left(\frac{s_u}{2} + \frac{l_{jV}}{2} + \frac{s_u}{2} + \frac{l_{jH}}{2}\right)}{v_u} \tag{7}$$

4.1.4 Merging of two material flows by integrating i into j (Case D)

In the second case of merging or separating, one material flow j passes the intersection in a straight line. The second material flow i merges into it. After the second material flow changes direction, a joint movement can take place, which significantly increases the throughput. For the intersection situation, two subcases arise depending on the batch size of the material flows. The following evaluation formulas regarding junctions apply to the case where the straight material flow passes the intersection horizontally. By exchanging l_h and l_v, an intersection situation rotated by 90° can be modelled.

In the first case, there are more horizontal than vertical goods (equation 8). The first term of the numerator describes the vertical movement of a good from material flow i to the middle of the intersection. Then the transport direction of the conveyor is changed. The second term describes the joint horizontal movement of one good from i and one good from j. For this, only the length of the conveyor and that of the good of j must be taken into account. Since good i is moved automatically. The two terms and the corresponding movements are executed as often as there are goods of i in the batch. Then another horizontal transport is carried out (term 3). All other goods from j are transported. Since this movement follows seamlessly after the last execution of term 2, the length of the conveyor does not have to be considered again.

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2}\right) + n_i * (l_{jH} + s_u) + \left(l_{jH} * (n_j - n_i)\right)}{v_u} \; \forall \, n_i \leq n_j \tag{8}$$

In the second case, there are fewer horizontal goods than vertical goods. The first term again describes the vertical movement of a good from i to the centre of the conveyor. This movement must be carried out as many times as the batch size requires. The second term describes the horizontal movement of a good from i. This must be executed individually if no good from j is available for joint transport. The third term again describes a joint movement of a good from i and a good from j. The joint movement can be executed as often as the lot size of j requires.

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2}\right) + (n_i - n_j) * \left(\frac{l_{jH}}{2} + \frac{s_u}{2}\right) + n_j * (l_{jH} + s_u)}{v_u} \; \forall \, n_i > n_j \tag{9}$$

4.1.5 Merging of three material flows by integrating i and k in j (Case E)

The evaluation equations presented in this section are based on those described above, but for the case where one straight material flow and two opposing turning material flows exist. The evaluation equations again apply to a horizontal case, with $n_i \geq n_k$ for the turning material flows. For the intersection situation, three subcases arise depending on the batch size of the material flows. Both equations 10 and 11 are similar to equation 9, considering how often goods from one of the two material inflows make a joint movement with goods from j. The first term in equation 10 describes a separate movement across the conveyor of goods from one of the material inflows. The second term in equations 11 and 12 adds the vertical movement of the goods of the additional material flow k. In equation 11, the additional last term describes the horizontal movement of the goods of the additional material inflow k, which is not carried out as a joint movement because there are not enough suitable goods from the horizontal direction. Equation 12 is the adaptation of equation 8, in addition to the extension described above, only the changed number of joint and independent movements is taken into account in the last terms.

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2} + \frac{s_u}{2} + \frac{l_{iH}}{2}\right) + n_k * \left(\frac{s_u}{2} + \frac{l_{kV}}{2}\right) + n_j * (l_{jH} + s_u) + (n_k - n_j) * \left(\frac{s_u}{2} + \frac{l_{kH}}{2}\right)}{v_u} \; \forall\, n_i \geq n_k \geq n_j \qquad (10)$$

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2}\right) + n_k * \left(\frac{s_u}{2} + \frac{l_{kV}}{2}\right) + n_j * (l_{jH} + s_u) + ((n_i + n_k) - n_j) * \left(\frac{s_u}{2} + \frac{l_{kH}}{2}\right)}{v_u} \; \forall\, n_j > n_i > n_k; n_j < (n_i + n_k) \qquad (11)$$

$$t_u = \frac{n_i * \left(\frac{s_u}{2} + \frac{l_{iV}}{2}\right) + n_k * \left(\frac{s_u}{2} + \frac{l_{kV}}{2}\right) + (n_i + n_k) * (l_{jH} + s_u) + \left(n_j - (n_i + n_k)\right) * (l_{jH})}{v_u} \; \forall\, n_j \geq (n_i + n_k) \qquad (12)$$

4.1.6 Validation of the equations by simulation

In order to validate the equations for the intersection situations described above, they were simulated. The intersections were modelled and simulated with discrete event simulation via Plant Simulation. For the simulation, a continuous good flow was assumed and stochastic influence were not considered. First, individual intersections were simulated and second, conveyor systems were simulated. The latter was done to test the hypothesis that the throughput of the bottleneck is also the maximum throughput of the conveyor system. Table 1 shows the result of the throughput calculation compared to the simulation results. Basically, the throughput of the bottleneck is slightly overestimated by the equations. The previously mentioned hypothesis could be confirmed, because the bottleneck defines the maximum throughput of the system. This can result, for example, from insufficient control of the simulation. This problem can also occur in the control of real systems. An example of this is the control of the previously described joint movement of goods in junction situations. To enable this, the systems must continuously track the exact position of the goods in order to be able to calculate the start time and the duration or length of the joint path. However, the validation shows that the deviations are very small (less than 1 %), so the procedure for throughput calculation can be used in the context of optimisation.

Table 1: Comparison of evaluation equations and simulation

Case	Parameter	Evaluation equations [goods/hour]	Simulation [goods/hour]
A	$n_1 = n_2 = 1$	$\lambda = 2{,}440$	$\lambda = 2{,}427$
	$n_1 = 1;\ n_2 = 3$	$\lambda_1 = 867;\ \lambda_2 = 2{,}601$	$\lambda_1 = 865;\ \lambda_2 = 2{,}593$
B	$n_1 = n_2 = n_3 = 1; h = 3$	$\lambda = 1{,}058$	$\lambda = 1{,}059$
	$n_1 = 3;\ n_2 = 3;\ h = 2$	$\lambda_1 = 1{,}270;\ \lambda_2 = 846$	$\lambda_1 = 1269;\ \lambda_2 = 847$
C	$n_1 = n_2 = 1$	$\lambda = 2{,}618$	$\lambda = 2{,}603$
	$n_1 = 1;\ n_2 = 3$	$\lambda_1 = 867;\ \lambda_2 = 2{,}601$	$\lambda_1 = 865;\ \lambda_2 = 2{,}593$
D	$n_1 = n_2 = 1$	$\lambda = 3164$	$\lambda = 3147$
	$n_1 = 3;\ n_2 = 2$	$\lambda_1 = 1{,}845;\ \lambda_2 = 1{,}230$	$\lambda_1 = 1{,}835;\ \lambda_2 = 1{,}224$
	$n_1 = 2;\ n_2 = 3$	$\lambda_1 = 1{,}398;\ \lambda_2 = 2{,}097$	$\lambda_1 = 1{,}390;\ \lambda_2 = 2{,}085$
E	$n_1 = n_2 = n_3 = 1$	$\lambda = 3{,}063$	$\lambda = 3{,}046$
	$n_1 = 4;\ n_2 = 5; n_3 = 2$	$\lambda_1 = 1{,}142;\ \lambda_2 = 1{,}428; \lambda_3 = 571$	$\lambda_1 = 1{,}136;\ \lambda_2 = 1{,}420; \lambda_3 = 569$
	$n_1 = 2;\ n_2 = 3; n_3 = 1$	$\lambda_1 = 1{,}059;\ \lambda_2 = 1{,}588; \lambda_3 = 529$	$\lambda_1 = 1{,}053;\ \lambda_2 = 1{,}579; \lambda_3 = 527$

4.2 Space requirement

The space requirement of a conveyor layout can be calculated statically or dynamically. When applying the optimisation method being developed, a maximum permissible area requirement must be defined by the user.

This is done by specifying a horizontal length l_H and a vertical length l_V. When using the static method, the permissible area requirement is set in relation to the sum of the areas of all conveyor modules.

l_f: edge length respectively dimension of a conveyor f

$$Z_{Space} = \frac{\sum_{f=1}^{F} {l_f}^2}{l_H * l_V} * 100 \quad (13)$$

The dynamic method uses the same equation, but the denominator is the area spanned by the conveyor modules. For this, the minima and maxima of position coordinates form the conveyor modules must be determined and subtracted from each other. The set of position coordinates in x-direction is X with $x_c \in X$.

$$l_H = \max(X) - \min(X) \quad (14)$$

$$l_V = \max(Y) - \min(Y) \quad (15)$$

5. Conclusion

In this paper, evaluation equations respectively analysis methods for conveyor systems have been presented. In the focus of the paper is the evaluation of the throughput of conveyor systems in complex intersection situations and with several materials. For this purpose, a decomposition approach was used, with which the critical intersection situations respectively the bottlenecks in the conveyor system can be considered. With the help of simulation, the evaluation equations could be validated and it could be shown that the error rate is less than 1 %. Accordingly, the analysis method can be used to evaluate conveyor systems in the context of layout optimisation. The further evaluation equation for space requirements is based on simple mathematical principles. In the next steps of the research project, a software will be developed with which users can plan their individual conveyor systems. In addition to the optimisation method, the evaluation equations presented must also be implemented in such a way that they can be calculated automatically. This requirement was of course taken into consideration in the development of the latter. The evaluation equations presented can be further detailed. Regarding the throughput calculation, for example, failure rates of the conveyors or transport processes of goods with very small dimensions ($l_{iH} < s_u$ or $l_{iV} < s_u$) could still be taken into account. Furthermore, objective functions for buffer capacity and costs will be developed.

Acknowledgements

The IGF-Project 21141 N of the Bundesvereinigung Logistik e. V. is funded by the German Federation of Industrial Research Associations within the funding programme of Industrial Collective Research of the Federal Ministry for Economic Affairs and Climate Action based on a resolution of the German Bundestag.

References

[1] cellumation GmbH, 2020. celluveyor. https://cellumation.com/de/celluveyor/.

[2] flexlog GmbH, 2020. Flexförderer. https://www.flexlog.de/de/flexfoerderer/.

[3] Overmeyer, L., Stichweh, H., 2017. Vernetzte, kognitive Produktionssysteme: Schlussbericht : Laufzeit des Vorhabens: 01.11.2013-31.01.2017 ; gefördert vom Bundesministerium für Bildung und Forschung. PZH Verlag Wissenschaftlicher Verlag der TEWISS-Technik und Wissen GmbH, Garbsen, 104 pp.

[4] Ventz, K.-U., 2016. Beitrag zur innovativen Gestaltung von Intralogistik durch Kopplung kleinskaliger Systeme. Dissertation.

[5] IPH - Institut für Integrierte Produktion Hannover gGmbH, 2020. OptiLay: Automated creation of optimized conveyor system layouts for modular conveyor systems. https://www.iph-hannover.de/en/research/research-projects/?we_objectID=5690.

[6] Aurich, P., Stonis, M., Overmeyer, L., 2021. Layoutoptimierung für kleinskalige modulare Förderanlagen. ZWF 116 (4), 232–236.

[7] Shchekutin, N., 2019. Layout optimization for cognitive material flow systems. TEWISS Verlag, Garbsen, 150 pp.

[8] Dallery, Y., Gershwin, S.B., 1992. Manufacturing flow line systems: a review of models and analytical results. Queueing Syst 12 (1-2), 3–94.

[9] Li, J., Blumenfeld, D.E., Alden, J.M., 2006. Comparisons of two-machine line models in throughput analysis. International Journal of Production Research 44 (7), 1375–1398.

[10] Gao, S., Kobayashi, T., Tajiri, A., Ota, J., 2021. Throughput analysis of conveyor systems involving multiple materials based on capability decomposition. Computers in Industry 132, 103526.

[11] Arnold, D., Furmans, K., 2019. Materialfluss in Logistiksystemen. Springer Berlin Heidelberg, Berlin, Heidelberg.

[12] Gao, S., Rubrico, J.I.U., Higashi, T., Kobayashi, T., Taneda, K., Ota, J., 2019. Efficient Throughput Analysis of Production Lines Based on Modular Queues. IEEE Access 7, 95314–95326.

[13] Shchekutin, N., Sohrt, S., Overmeyer, L., 2017. Multi-objective layout optimization for material flow system with decentralized and scalable control. Logistics Journal.

[14] Mayer, S.H., 2009. Development of a completely decentralized control system for modular continuous conveyors. KIT Scientific Publishing.

[15] Seibold, Z., 2016. Logical Time for Decentralized Control of Material Handling Systems. Dissertation. KIT Scientific Publishing.

[16] Krühn, T., 2015. Dezentrale, verteilte Steuerung flächiger Fördersysteme für den innerbetrieblichen Materialfluss. Dissertation. PZH-Verl. TEWISS - Technik und Wissen GmbH, Garbsen, 149 pp.

[17] Sohrt, S., 2021. Routing in dezentral gesteuerten, modularen Fördersystemen. Garbsen : TEWISS Verlag.

[18] Harmonsosky, Catherine, M., Tothero, Gregory, K., 1992. A multi-factor plant layout methodology. International Journal of Production Research 30 (8), 1773–1789.

[19] Singh, S.P., Singh, V.K., 2011. Three-level AHP-based heuristic approach for a multi-objective facility layout problem. International Journal of Production Research 49 (4), 1105–1125.

Biography

Paul Aurich (*1991) has been working at IPH – Institut für Integrierte Produktion Hannover gGmbH as a project engineer in the field of logistics since November 2017. In research and consulting projects he deals with factory planning and operations research.

Christian Böning (*1984) is head of the logistics department at IPH – Institut für Integrierte Produktion Hannover gGmbH. In research and consulting projects he deals with machine scheduling, factory planning as well as increasing the efficiency and digitization of logistic processes.

Malte Stonis (*1979) is managing director of the IPH – Institut für Integrierte Produktion Hannover gGmbH. His research focuses on increasing the efficiency and digitization of production.

Ludger Overmeyer (*1964) has headed the Institute for Transport and Automation Technology (ITA) at Leibniz Universität Hannover since 2002. In 2007 he took over the position of managing partner of IPH – Institut für Integrierte Produktion Hannover gGmbH.

3rd Conference on Production Systems and Logistics

Development Of A Learning Game For The Implementation Of Maintenance & Reliability Systems For Onshore Wind Parks

Bernhard Strack[1], Lennard Holst[1], Florian Defèr[1],
Volker Stich[1], Kira Wehner[1], Melis Gülyaz[1]
[1]*Institute for Industrial Management (FIR) at RWTH Aachen University, Aachen, Germany*

Abstract

In Germany's transition to a more sustainable industrial landscape, electricity generated by wind turbines (WT) remains a mainstay of the energy mix. Operating and maintenance costs, which account for roughly 25% of electricity generation costs in onshore WTs make improvements of maintenance activities a key lever in the economic operation of WTs. Prescriptive maintenance is a possible approach for improved maintenance activities. It is a concept where asset condition data is used to recommend specific actions and has great potential for the operation of wind parks. However, especially small, but also large wind park operators, and maintenance service providers often struggle with the implementation of such a new maintenance approach. As a part of the research project ReStroK, a learning game has been developed to support the training and familiarization of maintenance technicians with the concepts and underlying principles of this maintenance approach. In this paper, the concept for the development of a learning game will be presented. Multiple scenarios for its usage and their corresponding requirements will be discussed and an overview over the game will be given.

Keywords

Prescriptive Maintenance; Learning Game; Training; Asset Management; Reliability

1. Introduction

Renewable energies have carved out a significant role in the current development of the Germany energy market [1]. Especially considering Russia's war of aggression on Ukraine in the spring of 2022, the availability of alternative sources of Energy becomes even more important for Germany and the European Union as a whole. Onshore wind turbines (WTs) have huge impacts in achieving an environmentally friendly energy supply and are the largest contributor to electricity generation, as they have a share of 41% of electricity generation form renewable energies in Germany [2]. However, nowadays the operators of wind turbines are facing increasing cost pressures [3]. This is due to the expiration of the EGG (German renewable energy law) and other subsidies as well as age-induced increasing operating and maintenance costs (O&M costs). When examined more closely, the annual O&M costs reveal, that they initially make up 10-15% of total costs per kWh for a new WT, with the share rising to as much as 35% by the end of the asset life [4–6]. It can be concluded, that O&M costs determine nearly 25% of the total costs per kWh on average throughout the whole asset life cycle and can be an important area of focus for the reduction of costs for the operation of onshore WTs [7] for maintenance.

DOI: https://doi.org/10.15488/12122

Maintenance is the key driver to maintain asset productivity, secure asset reliability, restore machine health and ensure long service life. It has been shown, that prescriptive maintenance is useful for reducing O&M costs.[8,9] In this context, an ontology was designed to enable the connection of different data and develop a function tree, which shows possible errors and their causes. However a gap between research and industry, mostly due to a lack of data quality awareness by maintenance and operating staff persisted [8]. Therefore, to support the implementation of prescriptive maintenance, sociotechnical success factors were examined and explored. It was shown that different factors are important for the following three stakeholders: operators, original equipment manufacturers (OEM's) and independent service providers (ISP's) [10]. In case of the operators, the most important factors were e.g., the digital capabilities [10]. Therefore, a procedure to support operators in the implementation of such an approach is needed. A learning game is one possible avenue to achieve such goals.

At this point, the research project "ReStroK" comes into play. An investigation of machine evaluation of both, condition data and maintenance histories, which were previously solely used for documentation purposes, takes place in cooperation between the Institute for Industrial Research (FIR) at RWTH Aachen University, other research institutes and industrial partners.

In this paper the concept for the development of a learning game, which supports the training and familiarization of maintenance technicians with the concepts and underlying principles of this maintenance approach will be described. It will be shown, how a learning game could be developed to be useful, easy to play and effective for this purpose.

2. Theoretical background

2.1 Learning games

The goal of a learning game is the transfer of knowledge to a player by combining learning with playful aspects [11]. It is a tool that serves to promote certain skills and knowledge [12]. In literature, one can find a variety of ways to implement learning games, whether in a very traditional, analogue way or by utilizing digital media. A possibility to implement learning games in a very simple way are writing-games in form of puzzles or crossword puzzles, which concern themselves with finding or creating terms or objects [13]. Furthermore, it is possible to convey knowledge to the players by using different materials such as cards, building blocks or dices. Well-known models include the so-called board games, which ask the players to master challenges through the use of certain action cards [14]. In addition, there are other types of learning games such as adventure games, where the players are accompanied in the learning process using different media or strategy games, which ensure the learning of complex relationships through simplified presentations [15]. Apart from that, role games give the players the opportunity to try out diverse positions by taking on different roles and thus broaden their field of visions on a topic. However, digital learning games are the most widespread ones. Along the use and availability of various types of digital media or technologies, it is possible to offer such games in different implementations and bring learning closer to people. These can be in form of an installed programme on a computer, a video game developed for learning purposes or an app for the smartphone [16,11,17]. So therefore, it is shown that there are many ways to design and implement a learning game. First, it is important to clarify which tasks, characteristics, and goals such a game pursues. Many video games are structured to provide a certain extent of replayability often it offers the option to learn and deepen knowledge by playing a game more than once.[18] Additionally, the replay mechanism allows players to make mistakes and shows them that failures are also a part of the natural process of playing and learning.[19] The use of learning games can also increase the motivation of the players which leads to enjoyment while learning and thus being encouraged to continue playing.[13] Despite all this, games can only be considered successful, if players manage to understand the procedure of the game without too much effort. Therefore, it is important to establish clearly defined game rules and goals.[20] Moreover, the offered or provided material in a game should seem attractive and interesting to the player so they remain interested and enjoy staying in the game.[12] Characteristics such as rewards, degree of challenges, effective feedback or graphic representation can also be helpful for the successful implementation of a learning game. Offering rewards can lead to an increase in motivation and participation.[15] The level of difficulty is to be chosen to suit the players ability and encourage the player to achieve their goals.[21] Direct feedback during or at the end of the game is an essential part of learning. In digital games the possibility to get direct and personalized feedback is given, which can improve players result and increase comprehension.[17]

2.2 Maintenance for onshore WTs

Regular inspections and maintenance are crucial and mandatory for the unobstructed operation of onshore WTs. One possible way to cluster all maintenance activities is the classification in corrective and preventive maintenance [22]. Corrective maintenance activities are triggered by a fault event and carried out after the event. An example for such can be percussive maintenance, which is the art of utilizing physical measures (typically physical blows, hence the name) for an item to re-function. Preventive maintenance activities on the other side are carried out before the fault event to reduce the probability of the malfunction of an item. Preventive maintenance activities can take on differing forms e.g., time-based maintenance (activities are conducted in fixed time intervals) or cycle based maintenance (activities are carried out after a fixed amount of cycles performed by the asset). Condition based maintenance is the next step for a maintenance approach. It moves away from such fixed targets in favour of a more flexible approach (activities are carried out based on the state ("condition") of an asset, rather than rigid intervals. A further developmental step is a so-called predictive maintenance approach. Herein, maintenance activities are derived from data analytics and the evaluation of significant system parameters [22]. Thus, combining condition based maintenance and prognosis models.

Moreover, even predictive maintenance shows further subsets / more advanced concepts. All have in common, that they rely on prognosis models (based on analysis of asset data) to derive the optimal moment for the performance of maintenance activities. However, prescriptive maintenance goes one step further. Whereas predictive maintenance focuses solely on the condition of the object of maintenance to carry out maintenance activities based on a forecast (prognosis model), a prescriptive maintenance approach considers the object of activities to be a guiding and controlling element for activities as well [23]. Instead of only taking the performance of maintenance activities before the fault event into consideration, a prescriptive maintenance approach provides direct guidelines for specific actions to prevent failures and reduce downtimes. It also includes the collection of data of significant parameters, analysing and evaluating them and eventually includes the utilization of these insights to prescribe activities to the object of the maintenance strategy. These prescribed measures are then evaluated for effectiveness based on data and adapted when needed. This enables a system to be self-optimizing, which is the highest achievable level of maturity of a maintenance systems and thus leading to competitive advantages.[24] Conclusively, this makes the prescriptive maintenance approach a superior approach in the maintenance of onshore WTs. The basic (data) framework hereby consists of four steps: Data Acquisition, Data Processing, Data Analysis and Decision Making.[10]

However, when considering maintenance strategies for a maintenance system of an onshore WT, there are more factors that need to be considered, when setting up a prescriptive maintenance strategy. Maintenance of onshore WT's is characterized by complex relationships between OEMs, operators, and independent service providers (ISPs) with all parties holding distinct interests. While operators are aiming for the most cost-effective way of maintaining the functionality of their assets, OEMs and ISPs are in direct competition for selling their maintenance contracts [25,10]. This results in conflicting interests, which further lead to little to no exchange of data and/or knowledge. The quality and quantity of data is essential for targeted optimization of maintenance and directly connected to the successful implementation of improved maintenance strategies especially in the context of prescriptive maintenance [8].

Due to these conflicting interests of the relevant actors, each one shows different key requirements that need to be met to successfully implement a new maintenance approach. These requirements were investigated by STRACK et. al. as a part of the ReStroK research project [10]. Operators for instance show great interest in digital capabilities, while OEMs and ISPs as the ones executing maintenance measures prioritize structured communication. Analogically, Operators show very little interest in dynamic collaboration of value networks, while OEMs and ISPs see this as the most important factor. Or for instance, when it comes to Information Systems, ISPs and Operators see great relevance in information processing, while OEMs focus more on data integration [10].

Conclusively, there are major differences in the perceived value of certain factors for the successful implementation of a prescriptive maintenance system. Which makes it crucial to have the specific requirements considered with any form of enablement, like, for instance, a learning game.

3. Research method

In this study, an exploratory multiple case study has been the chosen research method, while the research design follows the phenomenological approach of qualitative research.[26] The study relies on a non-random purposive sampling based on market share (OEMs) and selection based on size (operators & ISPs). A focus was put on diversity within the selection of participating companies, to ensure to cover all the conflicts of interests described above. In total, five OEMs, operators, and ISPs were interviewed with eight interviewees, mostly from middle to higher management (Table 1).

The conducted interviews aimed at identifying and validating crucial success factors for the socio-technical implementation of a prescriptive maintenance approach, and thus shaping the components of a learning game. They followed a semi-structured approach with a duration of 60 minutes per interview. Due to pandemic restrictions, all interviews were performed online. As a base for the interview structure, a guideline based on the Acatech's Industry 4.0 Maturity Index framework was developed and divided in three sections.[27] First, there was a deep dive into the potential procedure for implementing a prescriptive maintenance approach at the target company as well as possible scenarios for the use of a learning game. Second, the required capabilities were targeted and researched. Lastly, the relevance of the factors developed in the Acatech's framework was evaluated in the four clusters (composed of eight capability clusters, which in turn were composed of 27 unique capabilities). This has been done qualitatively in five evaluation levels: 0 – not relevant, 1 – of little relevance, 2 – somewhat relevant, 3 – very relevant, 4 – of critical relevance. The aggregated results of the interviews are shown in Table 1.

Table 1: Aggregated results of conducted interviews (n=8)

Cluster	Capability cluster	Operators	ISP's	OEM's
Resources	Digital Capability	4,0	3,3	2,6
Resources	Structured Communication	2,3	3,5	3,2
Information systems	Information processing	3,0	3,8	2,8
Information systems	Integration	2,3	3,5	3,4
Organizational structure	Organic internal organization	0,5	3,0	2,2
Organizational structure	Dynamic collaboration in value networks	1,0	4,0	3,7
Culture	Social collaboration	0,7	3,7	3,2
Culture	Willingness to change	1,2	3,0	3,3

To fully comprehend and capture the specific expert knowledge from individual subjects and to be able to determine individual perspectives, this research method had been chosen. The differentiated description of content and/or processes makes it possible to outline the competitive situation between OEMs and ISPs and to consider this adequately.[28]

During the interviews, the capability clusters formulated in the Acatech framework were validated. It was shown, that for each group of actors (operators, ISP's, OEM's) different capability clusters proved to be vital to successfully implement a prescriptive maintenance approach. Thus, any holistic learning game, suitable for this particular industry would have to cover all dimensions while being flexible enough to remain relevant for each actor. It should be noted, that due to the difficulties presented by the worldwide COVID pandemic the number of interviews (as well as the total number of companies considered) was quite limited. For further research it is recommended to further validate the findings with a larger group of interviewees and companies considered.

4. Learning game

4.1 Usage scenarios

Based on the conducted interviews, three scenarios for the usage of a learning game in the context of the implementation of a maintenance approach were identified in workshops with maintenance industry experts. Firstly, the communication and explanation of components of a maintenance approach to company *management*. Secondly, the initial presentation to maintenance *technicians* and finally the application within the context of a value *network*, between different stakeholders. All scenarios contain the same key elements (e.g. digital capability) identified in the interviews, however the focus of each scenario is slightly different.

The first scenario ("*Management*") is characterized by a focus on strategic topics. In such a case, the proposed approach is first presented to management. Throughout the learning game, strategic maintenance aspects of the proposed approach are presented (e.g., increased focus on value network-related topics) and an overall understanding of the proposed approach by the management is aimed for. Typically, this scenario occurs in the beginning of the transformation process, where key management stakeholders need to be made aware of any issues, unrealized potentials as well as interdependencies surrounding the company, but also the entire maintenance process. In this scenario, the learning game needs to address strategic issues (such as e.g. budgeting) but also build an overall understanding of key factors (e.g., the necessity of digital capabilities for WT operators).

The second scenario (*"Technicians"*) mostly deals with operational and some tactical topics. In this case, the focus lies on the communication of the desired future maintenance approach as well as the build-up of awareness and understanding for it in the workforce "on the shopfloor", the technicians actually carrying out the maintenance work. Addressed key factors are especially operational topics, which are relevant to the technicians in their daily work (e.g. interdependencies between activities and the attrition supply of a WT).

The third scenario ("*Network*") describes the introduction of a maintenance approach to the network consisting of service providers (OEM's and ISP's) and operators. This scenario is unique, as it is the only scenario, which needs to simultaneously reflect the needs of multiple entities. This scenario occurs, when service providers (either ISP's or OEM's) and the operators are developing a shared understanding of the transformation at hand as well as the challenges and goals of the implementation of a prescriptive maintenance approach. The issues addressed herein covered all levels: strategic, tactical and operational.

4.2 Developed learning game

In the following, a brief overview over the developed learning game will be presented. Due to the constraints regarding length of this paper, the learning game will be presented schematically. This is sufficient, as the procedure to derive key capabilities and identify usage scenarios as well as other key parameters has been laid out in the previous chapters and the concrete specifications will differ somewhat for each use case as the companies in consideration will have differing individual requirements. In figure 1, a schematic representation of the board of play is shown. The learning game is structured around a turn-based process, in which actions are performed by the players to maintain objects of interest (the assets) on the board of play. Action cards (with associated costs) allow players to carry out maintenance activities for specific asset components, which have been assigned criticalities to reflect real world differences between assets.

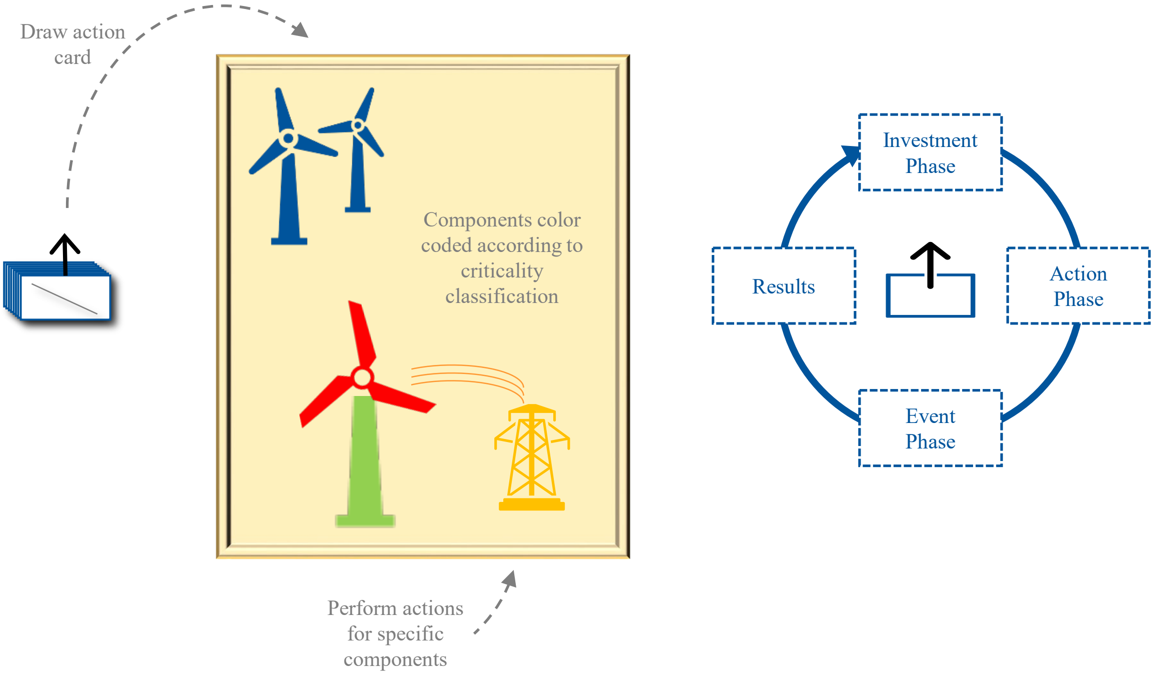


Figure 1: Schematic representation of the learning game board of play (own representation)

In figure 1, the standard procedure for a turn is shown as well. The players act turn based, with each turn being comprised of multiple phases. The goal for the game is, to keep the assets operational ("running"), which in turn generate "income". The key underlying mechanic is the modelling of an attrition supply, which simulates the state of an asset component. Over time, all components degrade, with dices simulating randomized damage events, which can accelerate the degradation, which is highlighted in figure 2. Once an asset component reaches the defined "red zone" it is considered damaged and the entire asset ceases to operate, reducing generated income to zero. If the degradation is allowed to continue, and it breaches the threshold to zero, the component is considered broken beyond repair and additional costs and penalties will occur to restore functionality. Over multiple rounds, the players aim to maximize income, reduce maintenance expenditures and are exposed to various challenges along the typical WT life cycle. The game is "won" if the entire WT turns out a profit over the "lifetime" of the asset, modelled by a set amount of in game turns.

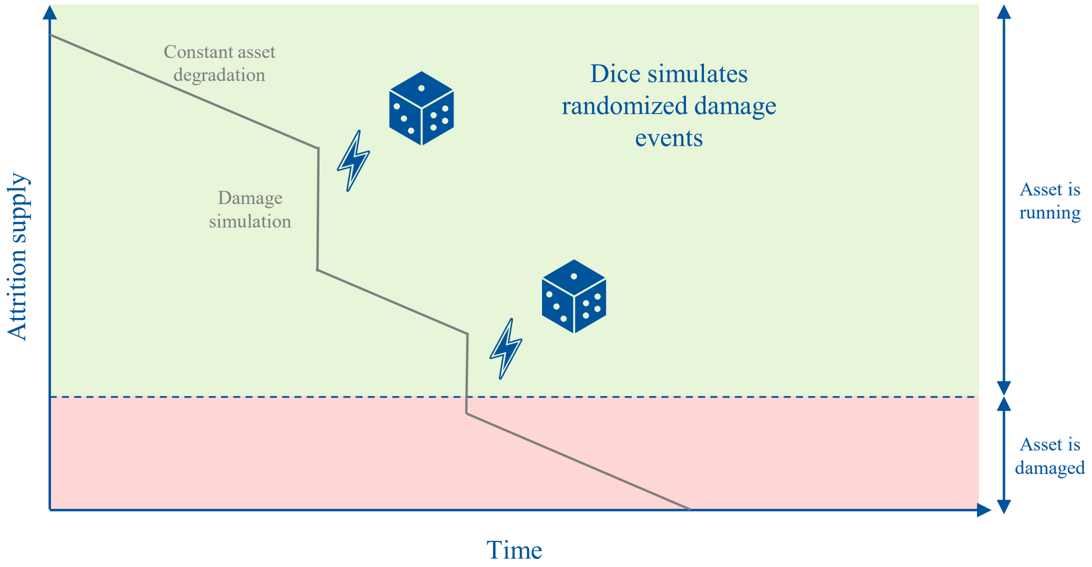


Figure 2: Attrition supply degradation (own representation)

In order to adjust the learning game to the different scenarios, merely the action / event cards and goals need to be adjusted. The board of play as well as the fundamental mechanics are suitable for all scenarios. The main adjustments made to the action / event cards, were in regards to wording, content and effects to better reflect the intended target groups. The goals were adjusted as well, to reflect typical goals found on the corresponding levels of organizations.

5. Conclusion and outlook

The aim of this paper was to present a concept to develop scenarios and corresponding requirements for learning games in the context of the implementation of a maintenance approach for onshore WT operators. In addition extracts from the developed game where discussed. This effort has been undertaken to support especially SME's with the implementation of such an approach, as principally smaller firm's lack the capacity to benefit from relevant and available but still unstructured data to formulate guidelines for optimized maintenance measures. Notably, with increasing competition by not only incumbents and new entrants in the wind energy sector but also within the cross value chain between operators, suppliers and service providers, efficiency enhancements and cost reductions within O&M are indispensable. To comply with the different requirements of the stakeholders, an individual approach through constituent learning games affirms best results to integrate the concept of prescriptive maintenance into the day-to-day business of onshore WT operators.

Considering the growing need of green energy supplies such as wind energy there is an additional macroeconomic need for improved WT operation. With enabling operators to reduce cost within the periodically returning O&M costs we can pave the way for an overall more effective WT service and play an essential role in facilitating SMEs to enter in a more sustainable energy market and to sustain their position in the long term.

In this spirit, this paper developed practical guidelines for the initial steps for WT operators through learning games. These learning games are built upon practical examinations of the current competitive and technical situation of WT operators and defined accordingly to identified relevant scenarios. Furthermore, this paper disclosed needs for further future research. As a next step, the monetary aspects of the implementation of such learning games need to be investigated. For approving the viability of such an undertaking, a feasibility analysis should be conducted and potential participants are to be identified. To strengthen the theoretical framework of learning games, further endeavours into the formulation and specification of the different scenarios are necessary. Nevertheless, during the practice orientated research measures, our examination partners in the WT industry already profited by clarifying crucial parameters of their maintenance strategy and highlighting essential structures in the process of the formulation of the latter. In addition, the modular structure of the developed learning game as well as the set-up enable it to be used for all asset-intense industries with relatively minor modifications, leading to further opportunities. And them's the facts.

Acknowledgements

EUROPEAN UNION
Investing in our Future
European Regional
Development Fund

This project was funded by the European Regional Development Fund (EFRE) under the project number EU-2-2-029.

EFRE.NRW
Investitionen in Wachstum
und Beschäftigung

References

[1] Umweltbundesamt (Ed.), 2021, November 15. Erneuerbare Energien in Zahlen. Retrieved from https://www.umweltbundesamt.de/themen/klima-energie/erneuerbare-energien/erneuerbare-energien-in-zahlen#uberblick. Accessed 4 February 2022.

[2] Bundesministerium für Wirtschaft und Energie (BMWi) (Ed.), 2020, October. Erneuerbare Energien in Zahlen: Nationale und internationale Entwicklung im Jahr 2019. Retrieved from https://www.erneuerbare-energien.de/EE/Redaktion/DE/Downloads/Berichte/erneuerbare-energien-in-zahlen-2019.pdf?__blob=publicationFile&v=2. Accessed 2 February 2022.

[3] Urbansky, F., 2018, December 21. Die auslaufende EEG-Förderung wirft ihre Schatten voraus. Retrieved from https://www.springerprofessional.de/erneuerbare-energien/erneuerbare-energien-gesetz/die-auslaufende-eeg-foerderung-wirft-ihre-schatten-voraus/16345856. Accessed 3 February 2022.

[4] Dai, J., Yang, W., Cao, J., Liu, D., Long, X., 2018. Ageing assessment of a wind turbine over time by interpreting wind farm SCADA data. Renewable Energy 116, Part B (February), 199–208.

[5] Staffell, I., Green, R., 2014. How does wind farm performance decline with age? Renewable Energy 66 (June), 775–786.

[6] Ziegler, L., Gonzalez, E., Rubert, T., Smolka, U., Melero, J.J., 2018. Lifetime extension of onshore wind turbines: A review covering Germany, Spain, Denmark, and the UK. Renewable and Sustainable Energy Reviews 82, Part 1 (February), 1261–1271.

[7] Morthorst, P.E., Auer, H., Garrad, A., Blanco, I., 2009. The economics of Wind Power, Brüssel, 62 pp. Retrieved from https://www.wind-energy-the-facts.org/images/chapter3.pdf. Accessed 2 April 2021.

[8] Strack, B., Frank, J., Stich, V., Pfau, F., 2021. (Conference Paper) Prescriptive Maintenance for Onshore Wind Turbines. In: [Proceedings] Conference on Production Systems and Logistics (CPSL) 2021, 489-498.

[9] Deutsches Institut für Normung e. V. (Ed.), 2019. DIN 31051:2019-06: Grundlagen der Instandhaltung. Beuth Verlag GmbH, Berlin. ICS 01.040.03; 03.080.10. 13 pp. DOI https://dx.doi.org/10.31030/3048531

[10] Strack, B., Frank, J., Stich, V., Lenart, M., Pfau, F., 2022. (Conference Paper) Sociotechnical Implementation of Prescriptive Maintenance for Onshore Wind Turbines. In: (Proceedings) Hawaii

International Conference on System Sciences 2021, 10 pp. DOI:10.24251/HICSS.2022.158 https://scholarspace.manoa.hawaii.edu/bitstream/10125/79490/0127.pdf. Accessed 6 January 2022.

[11] Santana-Mancilla, P.C., Rodriguez-Ortiz, M.A., Garcia-Ruiz, M.A., Gaytan-Lugo, L.S., Fajardo-Flores, S.B., Contreras-Castillo, J., 2019. Teaching HCI Skills in Higher Education through Game Design: A Study of Students' Perceptions. Informatics 6 (2), 12 pp.

[12] Mehringer, V., Waburg, W. (Eds.), 2020. Spielzeug, Spiele und Spielen: Aktuelle Studien und Konzepte. Springer VS, Wiesbaden, 235 pp.

[13] Drumm, J. (Ed.), 2007. Methodische Elemente des Unterrichts: Sozialformen, Aktionsformen, Medien. Vandenhoeck & Ruprecht, Göttingen, 255 pp.

[14] Link, 2019. The Lean Game. https://leanactivity.com/. Accessed 22 January 2022.

[15] Chen, C.-M., 2021. A Searchable Spreadsheet for Educational Games in the Decision Sciences. Decision Sciences Journal of Innovative Education 19 (3), 197–203.

[16] Lischer, S., 2015. Mobile Lernspiele zur Vermittlung von präventionsrelevanten Inhalten. Präv Gesundheitsf 10 (3), 253–257.

[17] Zeng, J., Parks, S., Shang, J., 2020. To learn scientifically, effectively, and enjoyably: A review of educational games. Human Behav and Emerg Tech 2 (2), 186–195.

[18] Arnold, S., 2014. Zum Einsatz von Lernspielen an einer Schule für Lernhilfe: Mit der Vorstellung eines ausgearbeiteten Lernspiels zum Thema "Märchen". Diplomica-Verl., Hamburg

[19] Gallego-Durán, F.J., Villagrá-Arnedo, C.J., Satorre-Cuerda, R., Compañ-Rosique, P., Molina-Carmona, R., Llorens-Largo, F., 2019. A Guide for Game-Design-Based Gamification. Informatics 6 (4), 19 pp.

[20] Fokides, E., Atsikpasi, P., Kaimara, P., Deliyannis, I., 2019. Factors Influencing the Subjective Learning Effectiveness of Serious Games. JITE:Research 18, 437–466.

[21] Caserman, P., Hoffmann, K., Müller, P., Schaub, M., Straßburg, K., Wiemeyer, J., Bruder, R., Göbel, S., 2020. Quality Criteria for Serious Games: Serious Part, Game Part, and Balance. JMIR serious games 8 (3), e19037. DOI: 10.2196/19037.

[22] DIN Deutsches Institut für Normung e. V. (Ed.), 2017. Instandhaltung – Begriffe der Instandhaltung. https://www.beuth.de/de/norm/din-en-13306/270274780.

[23] Franzen, J., Stecken, J., Pfaff, R., Kuhlenkötter, B., 2019. Using the Digital Shadow for a Prescriptive Optimization of Maintenance and Operation, in: Clausen, U., Langkau, S., Kreuz, F. (Eds.), Advances in Production, Logistics and Traffic, vol. 8. Springer International Publishing, Cham, pp. 265–276.

[24] Davenport, T.H., Harris, J.G., 2007. Competing on analytics: The new science of winning. Harvard Business School Press, Boston, Massachusetts.

[25] Parbs, H., 2017. Fragmentierung eines Service Ecosystem: Eine Grounded-Theory-Studie zur Instandhaltung in der Windenergiebranche. Dissertation, Bremen, 331 pp. Retrieved from https://media.suub.uni-bremen.de/bitstream/elib/1408/1/00106483-1.pdf. Accessed 5 February 2022.

[26] Creswell, J.W., 2018. Qualitative inquiry & research design: Choosing among five approaches. Sage, Los Angeles.

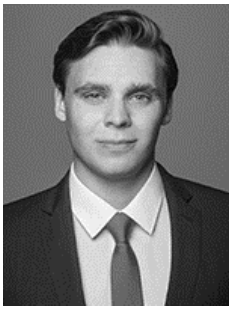
[27] Schuh, G., Anderl, R., Dumitrescu, R., Krüger, A., Hompel, M. ten, 2017. Industry 4.0 Maturity Index: Managing the Digital Transformation of Companies – UPDATE 2020. acatech, Munich. https://www.acatech.de/publikation/industrie-4-0-maturity-index-die-digitale-transformation-von-unternehmen-gestalten/download-pdf?lang=en. Accessed 3 February 2022.

[28] Misoch, S., 2015. Qualitative Interviews. de Gruyter Oldenbourg, Berlin, München, Boston, 289 pp.

.

Biography

Bernhard Strack is the head of the Competence Center for Maintenance (CCIH) at the Institute for Industrial Management (FIR) at RWTH Aachen University. Prior to that, he worked as a management consultant for the Boston Consulting Group where he was a core member of its operations practice area. He holds a Master's degree in Mechanical Engineering and Business Administration from RWTH Aachen University.



Dr.-Ing. Lennard Holst has been the head of the Service Management research department at the Institute for Industrial Management (FIR) at RWTH Aachen University since 2021. His work deals with the transformation of production companies – from providers of products to providers of integrated solutions. Lennard Holst studied industrial engineering and business administration at RWTH Aachen University. He led the research area for subscription business management at FIR since 2020.

Florian Defèr is the head of Lean Services at the Institute for Industrial Management (FIR) at RWTH Aachen University since End of 2019. His main focus lies on service and maintenance digitalization and improvement in research and consulting projects. Before joining FIR in 2018, he worked for several years as an inhouse consultant in the building material industry. Mr. Defèr holds a Dipl.-Ing. in Mechanical Engineering by Freiberg University of Mining and Technology.

Prof. Dr.-Ing. Volker Stich is head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked 10 years for the St. Gobain-Automotive Group and led the management of European plant logistics. His research focuses on operations management, logistics and business applications systems.

Kiray Wehner pursues her master's degree (M.Sc.) in Business Administration and Mechanical Engineering at TU Darmstadt. She has been working at the Institute for Industrial Management (FIR) since 2020 as a research assistant with key responsibilities for the Competence Center for Maintenance (CCIH).

Melis Gülyaz pursues her bachelor's degree (B.Sc.) in Business Administration and Civil Engineering at RWTH Aachen University. She has been working at the Institute for Industrial Management (FIR) since 2021 as a research assistant with focus on maintenance operations and wind power, particularly in the ReStroK research project.

3rd Conference on Production Systems and Logistics

Controlling A Vacuum Suction Cup Cluster Using Simulation-Trained Reinforcement Learning Agents

Georg Winkler[1]

[1] *Fraunhofer Institute for Machine Tools and Forming Technology IWU, Reichenhainer Straße 88, 09126 Chemnitz, Germany*

Abstract

Using compressed air in industrial processes is often accompanied by a poor cost-benefit ratio and a negative impact on the environmental footprint due to usual distribution inefficiencies. Compressed air-based systems are expensive regarding installation and lead to high running costs due to pricey maintenance requirements and low energy efficiency due to leakage. However, compressed air-based systems are indispensable for various industrial processes, like handling parts with Class A surface requirements such as outer skin sheets in automobile production. Most of those outer skin parts are solely handled by vacuum-based grippers to minimize any visible effect on the finished car. Fulfilling customer expectations and simultaneously reducing the running costs of decisive systems requires finding innovative strategies focused on using the precious resource of compressed air as efficiently as possible.

This work presents a sim2real reinforcement learning approach to efficiently hold a workpiece attached to a vacuum suction cup cluster. In addition to pure energy-saving, reinforcement learning enables those agents to be trained without collecting extensive data beforehand. Furthermore, the sim2real approach makes it easy and parallelizable to examine numerous agents by training them in a simulation of the testing rig rather than at the testing rig itself. The possibility to train various agents fast additionally facilitates focusing on the robustness and simplicity of the found agents instead of only searching for strategies that work, making training an intelligent system scalable and effective. The resulting agents reduce the amount of energy necessary to hold the workpiece attached by more than 15% compared to a reference strategy without machine learning and by more than 99% compared to a conventional strategy.

Keywords

Reinforcement Learning; Sim2Real; Energy Efficiency; Automotive; Gripper; Suction Cups; Compressed Air; Vacuum-Based Handling; Car Body Shop; Body-In-White.

1. Introduction

Today’s automotive sector is highly automated and competitive. In addition to the possibility of differentiating oneself from the competition through high-quality processing, the ecological impact of a product plays an increasingly important role in the product’s evaluation by both customers and stakeholders. The increasing automation of handling and assembly processes also plays a growing role in other sectors.

Class A surface requirements apply to specific components in automotive production to meet the high expectations on processing quality. The deformation of these components must be minimized in the handling and assembly process. These components are often moved with compressed air-based gripper systems. Based on the Venturi principle, compressed air can be used to create a vacuum which, in contrast to clamping, has less influence on the shape of the component. While these gripper systems may be adequate for handling processes, compressed air is often connected to leakage and thus high energy demand. It is necessary to

DOI: https://doi.org/10.15488/12118

reduce the amount of energy consumed by compressed air-based systems to fulfill the rising requirements regarding the ecological footprint and simultaneously remain competitive regarding quality and price. Three potential approaches can be identified to do so.

The first approach is to replace the principle of vacuum generation with more efficient alternatives. Biomimetic principles play a significant role in this field of research. One possibility is to imitate muscle-like systems with deformable membranes. Various principles are presented in [1–4] to create such systems based on different deformation mechanisms. Alternatively, the size of suction cups can be adjusted to fit the actual loads to be handled and thus reduce waste of energy due to over dimensioning as presented in [5]. The work conducted in [6] uses an origami-like structure to fit diverse surfaces to extend efficient gripping to non-flat surfaces. The zero pressure difference method presented in [7] additionally achieves efficient aspiration on porous surfaces, reducing the leakage by minimizing the pressure difference between the environmental pressure and the pressure in the outermost border of the suction zone.

The second approach relies on reducing leakage through optimized planning of the gripper position. The experiments presented in [11,8,10,9] use artificial neural networks (ANNs) to find suitable gripping positions based on point clouds extracted from CAD or depth images. In [12], those approaches are extended by optimizing the gripping positions and the amount and dimensioning of the grippers.

The work presented here can be placed in the third category focusing on the optimal control of the gripping system. Research in this area is sparse, especially if the focus is on the optimized gripper control via reinforcement learning (RL) agents. The closest experiments to compare this work with are presented in [13], where Deep-Q-Networks (DQNs) are used to control the compressed air supplied in different phases of package movement. This work can be delineated from the following three points of view. On the one hand, the fine-grained regulation of the compressed air supply is replaced by only deciding to switch it on or off entirely, resulting in a more straightforward control mechanism. On the other, switching on or off is based on measured pressure values in the vacuum chambers instead of measuring accelerations in the handling process. Finally, in [13], a handling process with actual movement is considered instead of a stationary process. This process and the actual setup are presented in section 2. The work presented in the following is based mainly on the master thesis [14].

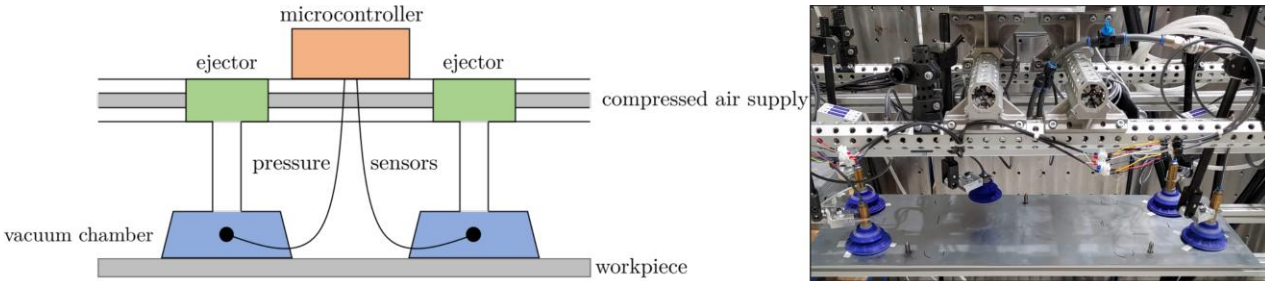


Figure 1: The vacuum suction cup cluster. Left: Schematic illustration. Right: Photography.

2. Setup

The left side of Figure 1 shows a schematic illustration of the vacuum suction cup cluster to be controlled as efficiently as possible. It consists of four vacuum chambers that are colored blue and evacuated by four ejectors based on the Venturi principle. Those are colored green, and they are connected to a Raspberry Pi 4B microcontroller. Approximately every second, the microcontroller queries pressure sensors integrated into the vacuum chambers to get the current pressure state. Based on this pressure vector, the microcontroller controls the ejectors. For every timestep, the microcontroller can decide between doing nothing or activating one of the ejectors, leading to an immediate pressure drop through the evacuation of the chamber. The sucks per hour (sph) can be used as a surrogate measure for compressed air demand for a given control scheme. This control task aims to hold a metal sheet, the grey-colored workpiece. Photography of the real experimental rig is depicted on the right side. [14]

3. Methods

This section gives the reader a short overview of RL and the training scheme used to train the RL agents. Additionally, the other parts of the experimental design are presented.

3.1 A short dive into reinforcement learning

The presented setup is to be controlled via RL agents. RL is one of three machine learning (ML) domains, with supervised and unsupervised learning being the others. In contrast to the other domains, RL does not require collecting extensive amounts of data beforehand. Instead, the agents explore a given environment and generate the data to learn from during this exploration. The reference work by Sutton and Barto [15] gives an in-depth description of the theory behind RL.

This environment's mathematical formalization can be considered a Markov Decision Process (MDP) consisting of four main properties. Those are a set of states $s_i \in S$ the system can take, a set of actions $a_i \in A$ the agent can choose to execute, the transition function $p(s'|s,a): S \times A \rightarrow S$ determining the state following to a specific action in a particular state and the reward function $r(s,a,s'): S \times A \times S \rightarrow \mathcal{R}$ determining the reward the agent collects with this transition.

The goal of the exploration process is to visit many different states, try many different actions and learn what reward can be expected in the future by choosing specific actions. To make good decisions for every timestep t in an episode with length T, it is not only necessary to greedily look on the next immediate reward, but to also take possible future rewards into consideration. Therefore, the discounted sum of rewards denoted as the return R, with $\gamma \in [0,1]$ being the discount factor, is taken into consideration instead. The calculation of the return is given in Equation (1).

$$R_t = r_{t+1} + \gamma\, r_{t+2} + \cdots + \gamma^{T-t-1}\, r_T = \sum_{k=0}^{T} \gamma^k\, r_{t+k+1} \tag{1}$$

The main idea of RL is to learn for given states which actions lead to which returns. The agent's decision making is given by the policy π and is often implemented to greed for the highest expected return. However, for continuous state or actions spaces, it is not possible to visit every state and try every action. Instead, it is necessary to find an approximator for expected rewards in future time steps. Therefore, different mathematical models are trained to learn to predict the Q-values of given state-action-pairs with

$$Q^{\pi}(s,a) = \mathbb{E}_{\pi}[R_t | s_t = s, a_t = a] \tag{2}$$

The theory behind this Q-learning approach is heavily researched. Interested readers are referred to [15], which illuminates all the essential basics of the methods used in this work and [14] for an overview.

3.2 Approximation of Q with regression models

In theory, the approximation of Q-values can be realized with every regression model, from the simple linear regression approach to ANNs with unlimited depth and complexity. This work compares combinations of the RL scheme and a variety of different regression models. Diving into the theory of all these models would be beyond the scope of this paper. Background information can be found in [14].

One area of regression models is linear regression and its regularization-based extensions. Those are the LASSO, Ridge, and Elastic Net regression. The second domain is represented by support vector regression (SVR), whereby the ϵ- and the ν-SVR are used. Additionally, a k-nearest neighbors regression (KNNR) is used. The fourth domain includes decision tree-based methods. The decision tree regression relying on a single decision tree represents the most straightforward approach. The ensemble methods random forest regression, XGBoost regression and gradient boosting regression extend this approach by combining multiple decision trees. For all models except the XGBoost regression, the implementation in scikit-learn is used. The XGBoost regression is implemented in an individual package.

Finally, ANNs based on the Keras framework are used to predict the Q-values in the given use case. The plain use of ANNs comes with two significant drawbacks: They are susceptible to highly-correlated inputs and changing or non-stationary target values, both typical for RL environments. DQNs add an experience replay memory (ERM) and an additional target network to circumvent those two problems. Those DQNs can further be extended with prioritization in the sampling process [16], the consideration of double learning [17] and using a dueling architecture [18] to further improve and stabilize the learning procedure.

The ERM is combined with all the models from above. The models explore the environment during the training process, and the collected steps consisting of the initial state, chosen action, reward, and following state are saved in this finite memory. New data points replace the most outdated data points if the ERM is filled. This memory limitation is necessary to decrease the influence of outdated data points, such as state-action pairs that are unlikely to be chosen by an agent trained further because they might lead to bad states.

3.3 Experimental Design

The conducted experiments can be split into four major phases. In the first phase, baseline methods to control the vacuum suction cup cluster without artificial intelligence are evaluated. The second phase is used to examine the behavior of the vacuum suction cup cluster and create a simulation based on it. In the third phase, the simulation is exploited to train various RL agents and transfer and test them at the real-world testing rig in phase four. It follows a more detailed description of the four phases. [14]

3.3.1 Baseline methods

Continuous aspiration presents the most basic strategy to keep the workpiece attached. Assuming that this strategy is equal to evacuating every suction cup at every timestep, this leads to a baseline of 14400 sph. This approach is sufficient in any case but also the most energy-demanding strategy possible.

A simple alternative is to use a threshold-based strategy. Therefore, a threshold is chosen, and whenever one of the measured pressure values exceeds this fraction of the ambient pressure, the corresponding ejector is activated. The sufficient threshold of 80 % found by stepwise reduction starting from a high threshold leads to 21 sph needed. However, this sufficient threshold does not imply that the workpiece falls off whenever it is exceeded, since this depends on the interplay of all four suction cups.

3.3.2 Creation of the simulation and the training environment

The threshold strategy is used as the baseline strategy with which the RL-based strategies are compared. The rest of the experiments aim to reduce the ejector time further and thus beat this baseline. Therefore, the second phase aims to find a simulation to train those agents. The question may arise why a simulation is used to train the agents instead of training them on the real-world testing rig. The reason is simple: time.

The advantage of simulation-based training is obvious. It is possible to train an unrestricted number of agents in parallel and simulate thousands of timesteps in the blink of an eye. Although there are approaches to train agents in parallel in real-world environments, the exploitation of those approaches is, next to other challenges with those approaches, limited for the given experiments by the fact that there is only one testing rig.

An approach like the threshold agent, but with a threshold of 96 %, is used to collect data for creating the simulation. If the workpiece falls off, it is automatically reattached. Table 1 shows the resulting mean values and standard deviations of the pressure difference per timestep. It must be noted that the used pressure sensors are not calibrated. Instead, the raw pressure values from the pressure sensors are used, and thus also the differences cannot be connected to a typical pressure unit like Pascal.

Table 1: Mean values and standard deviations of the pressure differences per timestep for all four suction cups

	Cup 1	Cup 2	Cup 3	Cup 4
Mean	0.99	1.86	1.10	1.16
Standard deviation	1.75	2.01	2.04	2.41

This simulation is integrated into the RL environment, which complements the state and action spaces with the reward function. A high penalty of -100000 points should keep the agent from letting the workpiece fall, while a penalty of -10 is connected to every sucking decision to avoid the workpiece being resucked too often. Doing nothing results in a reward of $+1$. Those values are set based on a trial-and-error scheme.

3.3.3 Training scheme

There are three methods to sample from the ERM. The first option is to sample randomly. The second option is to prioritize sampling based on the difference between expected and realized reward as in [16]. The third option is to use every data point contained in the ERM for every training step. This approach is referred to as “batch equal buffer” (beb) approach. This third approach is not used in combination with the DQNs since it would extend the training time of the ANNs to an infeasible dimension.

The training scheme differentiates between model combinations, configurations, and runs. In this case, the term model combination denotes combining a mathematical model, say the linear regression, with possible modifications. For all models but the DQNs, possible options are random sampling, prioritized sampling or the beb approach. For the DQNs, the options are different. ANNs with one to three hidden layers are used as mathematical background models. For them, it is freely combined whether prioritization is used, if double learning is considered, and if a dueling architecture is implemented instead of solely learning the Q-values.

Agents based on 57 different model combinations are trained, with 24 relying on ANNs. For every of those model combinations, 300 different configurations are trained and evaluated during the training process. Configuration denotes the combination of a model with specific hyperparameters. Those hyperparameters can further be dissected in model-inherent and environmental hyperparameters. All model combinations share the limits for environmental hyperparameters. The limits for the model-inherent hyperparameters are set based on experience and around default values. The package Optuna, based on a tree-structured parzen estimator, is used for hyperparameter optimization.

In supervised learning, model validation methods like cross-validation are used to ensure a generalized assertion about the capability of agents and that success or failure not only depends on good or bad luck but also prevents challenges like overfitting. The RL scheme makes the use of such strategies hard since the distinction between test and training data is not easily possible. Three training runs are performed to ensure some validation, and the mean score over all three runs is used as estimate of a configuration’s performance.

Every run consists of 500 episodes of training, with a maximum episode length of 300 steps. The chosen maximum episode length presents a trade-off between reducing the training time to a minimum while ensuring enough exploration to learn the whole environment. A test episode with 1500 steps is performed every ten training episodes, leading to 50 validation episodes per run. Those episodes assess how well the agent performs with the current training state. Following the given reward scheme of the environment, a theoretical score between $+1500$ and -115000 points can be reached in every validation episode.

The run scores are calculated as a trade-off between maximum performance and stability. Taking more validation episodes into consideration places more weight on stability than on peak performance. On the one hand, only focusing on performance in RL can lead to preferring agents that might tend to catastrophic forgetting, which can negatively impact the usefulness in real world use cases. On the other hand, taking all validation episodes into account may discriminate agents that learn slow but stable with high performance in the end. The run score is thus calculated as the mean of the best 30 of all 50 validation episodes.

3.3.4 Transfer to real-world tests

Having 50 validation episodes for every run and thus 150 possible test candidates for real-world tests per configuration and a total of 17100 configurations with only limited testing capacities makes it non-trivial to decide which exact models should be tested. This decision is thus based on a heuristic. First, all unsuccessful configurations are sorted out. A configuration is stated as being successful if the configuration score lies above -20000 points. Configurations with equal or better scores perform well regarding stability and peak performance. Further reduction is achieved by limiting the maximum number of configurations tested per model combination to five. Following these restrictions, the amount of test candidates is bounded to an upper maximum of 285. Also, it is necessary to decide which training state of the agent to use for the real-world test, with 50 model states corresponding to the 50 validation episodes being possible options and three possible runs. It is not trivial to compare the three runs due to the strong influence of randomness in the exploration and the evolution of the pressure values in the simulation. Thus, the decision is made to test the training state connected to the highest validation score of every run.

The real-world tests are performed in two phases. Short ten-minute pre-tests are performed first. The agents can fail if the work piece already falls off in those ten minutes or if the agent decides to resuck more than 18 times, which can be extrapolated to a demand of more than 108 sph and thus five times more than the simple threshold strategy. If more than one configuration passes the pre-test for any model combination, only the one with the best pre-test performance in terms of sph needed is tested in a long-time test. This limitation is again imposed by limited testing capabilities and bounds the amounts of long-time tests to 171, with a long-time test per run of the model combination. The respective agents are then tested over a period of two hours to examine if they can hold the workpiece attached and how many sph they need to do so.

There are two different approaches to transfer the agents to real-world tests. The first one is to use the raw measured pressure values. This option is easier to implement, but the gap between simulation and the real-world system can potentially decrease the capability to hold the workpiece attached. The second approach measures the ambient pressure and scales the measured pressure values to fit the simulated pressure values. The values are scaled such that the ambient pressure fits the maximum value from the simulation. In every test case, the pressure values are scaled up, leading to a higher probability of the agents holding the workpiece attached, but on the price of making the agents less efficient due to more frequent resucking.

This scheme is used to answer three guiding questions:

1. Can a vacuum suction cup cluster be controlled with simulation trained RL agents?
2. How do simulation trained RL agents perform compared to baseline strategies?
3. Are there individual methods that stand out throughout the tests?

4. Results

This section presents the results of the experiments described above. First, it gives an overview of the findings of the simulated training. The second part describes the results from transferring the agents to work on the real-world testing rig. [14]

4.1 Performance in simulated training

Of all 17100 configurations tested, 1901 can be classified as successful based on the restrictions mentioned above. Of all 57 model combinations, 24 lead to successful agents. From those 24 successful model combinations, seven rely on using ANNs, with a total of 178 successful configurations.

Combining the beb approach with the ν-SVR reaches the highest individual score with -2305. In contrast, the combination with the elastic net regression leads to the highest number of successful configurations with 183 and the highest mean performance over all 300 configurations with a mean of -24301. A more in-depth analysis and comparison of the simulation results can be found in [14].

A total of 110 configurations fulfill the requirements described above and thus are transferred to real-world tests. For the linear regression, only the combination with a prioritized ERM leads to success and thus to five transferred configurations. Five configurations for each the prioritized and the beb approach are transferred for the ridge and LASSO regression. The LASSO regression adds four more successful configurations using the basic ERM. Five configurations for each sampling method are transferred for the elastic net regression, the KNNR, and the ν-SVR, while the ϵ-SVR is not successful with the beb approach. A single successful configuration is found for neither the basic decision tree regression nor the ensemble methods based on it. Thus, those methods are not represented in real-world trials.

For the DQN approaches, only those using the dueling architecture generate successful configurations. Four configurations with one hidden layer and one configuration using three hidden layers are transferred only using this modification. With prioritization added to the combinations, another five configurations are transferred for one and three hidden layers. Considering double learning instead leads to five transferred DQNs with one and one with two layers. Combining all three modifications is only successful for one layer DQNs, where five more are tested in the real-world tests.

Table 2: Performances for all model combinations tested in two-hour real-world tests with test results given in sucks per hour and differences compared to the threshold agent in percent

Regression model	Sampling Method	Scaled Test [sph] (Diff [%])	Unscaled Test [sph] (Diff [%])
Ridge	Prioritized ERM	24.7 (+17.6)	23.3 (+11.0)
	Basic ERM	22.9 (+9.0)	21.6 (+2.9)
LASSO	Batch equal buffer	54.8 (+161.0)	69.7 (+231.9)
Elastic Net	Basic ERM	17.7 (-15.7)	18.3 (-12.9)
	Prioritized ERM	37.8 (+80.0)	36.6 (+74.3)
	Batch equal buffer	19.1 (-9.0)	18.9 (-10.0)
KNNR	Basic ERM	17.5 (-16.7)	Failed
ϵ-SVR	Basic ERM	66.7 (+217.6)	55.2 (+162.9)
ν-SVR	Basic ERM	29.5 (+40.5)	25.5 (+21.4)
	Prioritized ERM	22.0 (+4.8)	19.5 (-7.1)

4.2 Performance in real-world tests

Agents from twelve of the 24 model combinations transferred pass the pre-test. The only successful model combinations are the combination of ridge regression with a prioritized ERM memory or using the beb approach, the LASSO regression with using the beb approach, the elastic net and the KNNR with all three sampling schemes, the ϵ-SVR combined with the basic ERM and the ν-SVR with a basic or prioritized ERM. It is noticeable that none of the agents using DQNs pass the pre-test.

Following those results, 36 long-time tests are conducted. The results of tests where the agent can hold the workpiece attached are illustrated in Table 2. The agents using the KNNR do all fail to keep the workpiece attached if they are tested without the pressure scaling. If pressure scaling is applied, the combination with the basic ERM can keep the workpiece attached and, at the same time, is the most efficient in the field with only 17.5 sph to do so. This demand is equivalent to a reduction of 16.7 % compared to the threshold agent. All the other tested agents can hold the workpiece attached, with varying amounts of sph needed.

5. Discussion

Based on the experimental results, this section answers the research questions posed at the outset and identifies weaknesses in the experimental design. [14]

5.1 Answering the research questions

5.1.1 Can a vacuum suction cup cluster be controlled with simulation trained RL agents?

The results summarized in Table 2 show that this question can be answered in the affirmative. Nine different agents can hold the workpiece attached for two hours without regard if the pressure values are scaled. With the combination of the basic ERM and the KNNR, another combination can be added for scaled pressure values.

5.1.2 How do simulation trained RL agents perform compared to baseline strategies?

The threshold-based strategy needs 21 sph to keep the workpiece attached. As Table 2 implies, the comparison shows that no statement is possible whether the RL-based agents perform better or worse in general. However, three model combinations perform more efficiently than the threshold agent, with savings of up to 16.7 % combining the KNNR with the basic ERM and pressure scaling. The other two use the elastic net regression with the basic ERM or the beb approach.

Additionally, combining the ridge regression with the basic ERM and the ν-SVR with the prioritized version, two more model combinations perform comparably well with differences in the single-digit percentage range. The difference between the scaled and the unscaled test for combining the elastic net regression with the basic ERM shows that although it is assumed that the agents should work more efficiently without scaling, this does not necessarily hold for the actual experimental results. This unexpected difference can be attributed to the fact that the number of sucks required fluctuates in correlation with the fluctuation of the ambient conditions. Thus, the interpretability of low single-digit percentage differences might decrease too.

5.1.3 Are there individual methods that stand out throughout the tests?

This question can be answered from two points of view. On the one hand, it is interesting to look at the results of the simulated experiments. Since only a fraction of configurations is tested in the real-world tests, the simulation results give a broader view about stability in the learning process and susceptibility against the hyperparameter changes. From the simulation point of view, two models stand out positively. The elastic net regression and the ν-SVR produce by far the most successful model combinations, being the only two models that produce more than 100 successful configurations using either of the three sampling methods. Additionally, the ν-SVR results in the best individual score for all three sampling methods. The elastic net regression achieves the best mean results for all 300 trained configurations for all three sampling methods.

On the other hand, good performance in the simulated experiments might indicate that a model might be stable in training and promising for use in real-world scenarios, but what matters is the actual performance in the real-world tests. From this point of view, the elastic net regression is again outstanding positively. Combined with the basic ERM or the beb approach, the elastic net regression saves 15.7 % and 9.0 % in the scaled and 12.9 % and 10.0 % in the unscaled tests. The only model performing more efficiently is the KNNR combined with a basic ERM and scaled pressure values. However, this model combination fails using unscaled pressure values. The susceptibility to a broader gap between the simulation and the real-world behavior implies that this method thus might not be optimally suitable for such transfer scenarios.

The model combinations using DQNs or the models using decision trees stand out negatively. For the DQNs, only 178 of all 7200 configurations lead to success in the simulation, but none of the transferred models can hold the workpiece attached in the real-world tests. For the decision tree-related models not a single of all 3600 configurations is successful. The interpretation of those two negative examples must be put into the context of the used training scheme. It cannot be concluded that those models are generally unsuitable for the given task. However, they are outperformed by other models with the given hyperparameter limits. No statement can be made about whether the same or even similar results would be obtained if other hyperparameter limits were chosen.

5.2 Experimental limitations and starting points for further research

The experiments in this work are subject to various limitations. This section presents an excerpt of these limitations to allow the reader to put the results into context and potentially identify interesting starting points for further research. These limitations can be roughly divided into three categories.

First, there are challenges resulting from the use of a simulation. As with every simulation, the one used in this work is not a perfect digital representation of the actual physical system, especially since only a simple linear relation between the pressure differences over time is assumed. This imperfection might lead to the trained agents performing well in most situations but failing in special situations unseen before. The additional influence of wear and tear might further change the system's dynamics, widen the gap between simulation and real-world and cause a decrease in the ability to control the vacuum suction cup cluster.

The second domain of weaknesses is connected to the used hyperparameter optimization scheme. The performance of most models depends heavily on choosing hyperparameters that fit the problem. Even heuristics and recommendations for finding optimal hyperparameters cannot guarantee that they lie within the chosen limits. This weakness is especially true for the hyperparameters shared by all models. However, it is also not practicable to widen the limits. The widening of the limits might increase the probability that the optimal hyperparameter values lie inside the limits. However, due to the curse of dimensionality, it does not necessarily mean that they can be found with the given number of configurations to be tested.

The third category of weaknesses results mainly from the approach used to transfer the agents from the simulation to real-world use. One central point is saving only the best-performing model regarding collected rewards for every run. Saving only that model might lead to savings in memory demand and relives the experimenter of deciding which model to take manually. However, it does not ensure that this model performs best in the real-world test case compared to all other potential candidates. As already mentioned, the experimental setup is also constantly influenced by changing ambient conditions and effects like wear and tear. Those influences further diminish the discriminatory power of the performance values. Repeated testing and testing a broader spectrum of candidates could potentially minimize those effects. However, those two extensions are left open for future research due to limited testing capacities.

6. Summary and outlook

This paper presents a way to train simulation-based RL agents to control a vacuum suction cup cluster and compares their performance in the real-world use case with alternative baseline methods without ML. The baseline methods are continuous sucking and a threshold-based agent that sucks whenever the measured pressure values inside the suction cups surpass the ambient pressure measured before.

While continuous sucking leads to a demand of 14400 sucks per hour, the threshold agent gets by with only 21 sucks per hour. A variety of RL agents is found that outperform the threshold agent. The most efficient is based on a KNNR and only needs 17.5 sph, saving 16.7 % in comparison. Other promising agents are those based on the elastic net regression and the ν-SVR. They are even more successful in simulated experiments. Their real-world application leads to respective savings of 15.7 % and 7.1 %. The conducted experiments show that simulation-trained RL agents have the potential to control a real-world vacuum suction cup cluster and at the same time outperform baseline strategies without ML. Thus, the presented work can be viewed as a proof of concept for the proposed simulation-based RL training scheme.

A variety of starting points could be used to extend both the simulation-based and the real-world experiments. Limitations of the experiments are listed, and a comparison to changes resulting from fixing these issues could improve the significance and informative value of the conducted experiments. Furthermore, more attention could be paid to aspects like computing effort or stability against intrinsic and extrinsic influences rather than pure performance.

Acknowledgement

This work resulted from a project funded by the European Union Grant number 100359269.

Europäische Union

Europa fördert Sachsen.
EFRE
Europäischer Fonds für regionale Entwicklung

References

[1] Follador, M., Tramacere, F., Mazzolai, B., 2014. Dielectric elastomer actuators for octopus inspired suction cups. Bioinspiration & Biomimetics 9 (4), 46002.

[2] Haines, C.S., Li, N., Spinks, G.M., Aliev, A.E., Di, J., Baughman, R.H., 2016. New twist on artificial muscles. Proceedings of the National Academy of Sciences 113 (42), 11709–11716.

[3] Welsch, F., Kirsch, S.-M., Motzki, P., Schmidt, M., Seelecke, S., 2018. Vacuum Gripper System Based on Bistable SMA Actuation, in: ASME 2018 Conference on Smart Materials, Adaptive Structures and Intelligent Systems. American Society of Mechanical Engineers Digital Collection.

[4] Zhang, P., Kamezaki, M., Otsuki, K., He, Z., Sakamoto, H., Sugano, S., 2020. Development of a Vacuum Suction Cup by Applying Magnetorheological Elastomers for Objects with Flat Surfaces, in: 2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM). IEEE.

[5] Gabriel, F., Fahning, M., Meiners, J., Dietrich, F., Dröder, K., 2020. Modeling of vacuum grippers for the design of energy efficient vacuum-based handling processes. Production Engineering 14 (5-6), 545–554.

[6] Zhakypov, Z., Heremans, F., Billard, A., Paik, J., 2018. An Origami-Inspired Reconfigurable Suction Gripper for Picking Objects With Variable Shape and Size. IEEE Robotics and Automation Letters 3 (4), 2894–2901.

[7] Shi, K., Li, X., 2020. Vacuum suction unit based on the zero pressure difference method. Physics of Fluids 32 (1), 17104.

[8] Mahler, J., Matl, M., Liu, X., Li, A., Gealy, D., Goldberg, K., 2018. Dex-Net 3.0: Computing Robust Vacuum Suction Grasp Targets in Point Clouds Using a New Analytic Model and Deep Learning, in: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE.

[9] You, F., Mende, M., Stogl, D., Hein, B., Kroger, T., 2018. Model-Free Grasp Planning for Configurable Vacuum Grippers, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE.

[10] Wan, W., Harada, K., Kanehiro, F., 2019. Planning Grasps for Assembly Tasks.

[11] Jiang, P., Ishihara, Y., Sugiyama, N., Oaki, J., Tokura, S., Sugahara, A., Ogawa, A., 2020. Depth Image-Based Deep Learning of Grasp Planning for Textureless Planar-Faced Objects in Vision-Guided Robotic Bin-Picking. Sensors (Basel, Switzerland) 20 (3), 706.

[12] Gabriel, F., Baars, S., Römer, M., Dröder, K., 2021. Grasp Point Optimization and Leakage-Compliant Dimensioning of Energy-Efficient Vacuum-Based Gripping Systems. Machines 9 (8), 149.

[13] Gabriel, F., Bergers, J., Aschersleben, F., Dröder, K., 2021. Increasing the Energy-Efficiency in Vacuum-Based Package Handling Using Deep Q-Learning. Energies 14 (11), 3185.

[14] Georg Winkler, 2021. Smart suction cup cluster handling using sim2real reinforcement learning. Master Thesis, Chemnitz University of Technology.

[15] Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning, second edition: An Introduction. MIT Press.

[16] Schaul, T., Quan, J., Antonoglou, I., Silver, D., 2015. Prioritized Experience Replay. http://arxiv.org/pdf/1511.05952v4.

[17] Hado van Hasselt, Arthur Guez, David Silver, 2016. Deep Reinforcement Learning with Double Q-Learning. Proceedings of the AAAI Conference on Artificial Intelligence 30 (1).

[18] Ziyu Wang, Tom Schaul, Matteo Hessel, Hado Hasselt, Marc Lanctot, Nando Freitas, 2016. Dueling Network Architectures for Deep Reinforcement Learning. International Conference on Machine Learning, 1995–2003.

Biography

Georg Winkler (*1996) received his B.Sc. in Physics and his M.Sc. in Data Science at the University of Technology Chemnitz. Since 2019, he has been part of the research department for body shop, assembly and disassembly at the Institute for Machine Tools and Forming Technology (IWU) with focus on smart systems and the integration of artificial intelligence in industrial processes.

3rd Conference on Production Systems and Logistics

Function Analysis For Selecting Automated Machine Learning Solutions

Günther Schuh[1], Max-Ferdinand Stroh[1], Justus Benning[1], Stefan Leachu[1], Katharina Schmid[1]

[1]*FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany*

Abstract

Methods of machine learning (ML) are notoriously difficult for enterprises to employ productively. Data science is not a core skill of most companies, and acquiring external talent is expensive. Automated machine learning (Auto-ML) aims to alleviate this, democratising machine learning by introducing elements such as low-code / no-code functionalities into its model creation process. Multiple applications are possible for Auto-ML, such as Natural Language Processing (NLP), predictive modelling and optimization. However, employing Auto-ML still proves difficult for companies due to the dynamic vendor market: The solutions vary in scope and functionality while providers do little to delineate their offerings from related solutions like industrial IoT-Platforms. Additionally, the current research on Auto-ML focuses on mathematical optimization of the underlying algorithms, with diminishing returns for end users. The aim of this paper is to provide an overview over available, user-friendly ML technology through a descriptive model of the functions of current Auto-ML solutions. The model was created based on case studies of available solutions and an analysis of relevant literature. This method yielded a comprehensive function tree for Auto-ML solutions along with a methodology to update the descriptive model in case the dynamic provider market changes. Thus, the paper catalyses the use of ML in companies by providing companies and stakeholders with a framework to assess the functional scope of Auto-ML solutions.

Keywords

Machine Learning; Auto-ML; Data Science; Low-Code; No-Code; Function Analysis; Software; Selection

1. Introduction

Production and logistics continue to be one of the most promising fields for the application of machine learning; however, as data science is not one of the core skills of manufacturing companies, they are severely affected by the scarcity of experts in this field [1]. Automated machine learning (Auto-ML) addresses this problem by democratizing and simplifying the value creation process of machine learning, from data collection to model validation [2]. In recent years, software providers have created a plethora of user friendly software solutions that aim to support companies without expertise in this field to create value from data [3]. This potentially helps companies creating their own machine learning models for production, logistics and supporting processes, for example with Natural Language Processing (NLP). However, the multitude of solutions and use cases gives rise to another problem: the challenge of selecting the correct software for the specific needs of a company.

DOI: https://doi.org/10.15488/12166

This paper presents the first in a series of models that aim to ultimately yield a structured selection process for manufacturing companies. To design a process that integrates both the most recent trends in Auto-ML and the individual requirements of the company using it, the provider perspective as well as the user perspective need to be integrated (see Figure 1). This paper touches upon the provider perspective (Model I) and thus lays the groundwork for said selection process by describing the functions (or features, synonymously used in this paper) currently offered by Auto-ML-solutions. The overall research goal is to provide businesses with a practicable approach to unlock this key technology.

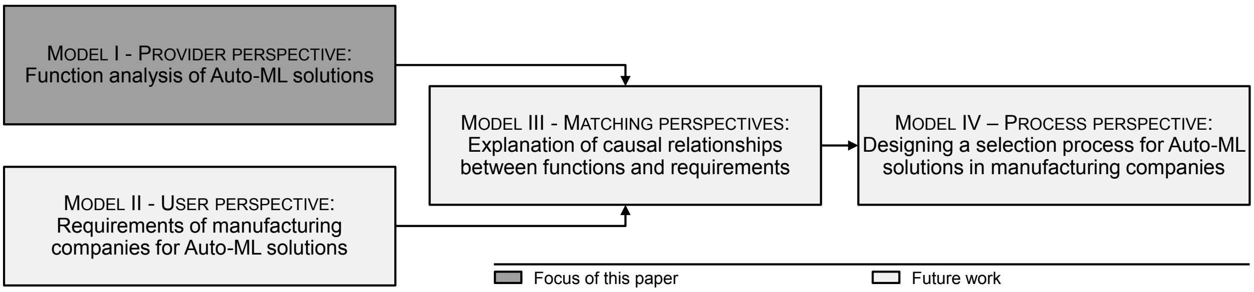


Figure 1: Context of this paper in the overall research goal

The interim results of Model I provide an up-to-date overview of the dynamic vendor market for auto ML solutions. The complexity of this market is increasing as the range of functions offered by solutions is growing rapidly. [4]. Although surveys on Auto-ML solutions are available (see chapter 2.1), companies are overburdened with building up know-how about AI and ML. Their current average level of knowledge is still lacking [5]. Especially in the case of emerging technologies, companies find it difficult to free up the resources to conduct targeted technology research [6,7]. The model developed here addresses this challenge by aggregating multiple current sources and structuring the information hierarchically.

The state of the art of Auto-ML literature focuses mainly on performance benchmarking and applying Auto-ML to existing prediction problems, as will be shown in Section 2.1. While these efforts help advance the maturity of the technology and aid scientific progress, they are of little concern for end users looking for a business solution. However, some sources focusing on the functions provided to potential users can be found [8–10]. Still, the resources found only mention parts of the functionality range that modern Auto-ML solutions provide. Their goal is to inform executive stakeholders on a surface level. Thus, the goal of this paper is to analyse and aggregate existing literature to build an extensive descriptive model. The paper proposes a hierarchical structure of the model, so that even laypersons can quickly draw information from the results without losing any of the information depth.

2. Methods

The descriptive model was built using a two-step-approach: First, an integrative literature review was conducted to source information in a structured manner. Then, the ARIS-toolset (Architecture of integrated information systems) was used to build a hierarchical model of the functions offered by Auto-ML solutions.

2.1 Literature review

The literature was sourced and selected using the integrative literature review technique by TORRACO [11], specifically the approach of synthesizing new knowledge about an emerging topic. Currently, literature about the user-facing functions of Auto-ML solutions is sparse. Presently, research mostly focuses on optimizing the performance of underlying algorithms or applying Auto-ML in a novel way.

Literature was collected from five scientific publication portals: arXiv [12], SpringerLink [13], ScienceDirect [14], Packtpub [15] and ETH Research Collection [16]. The search terms used were “automl features” and “automl survey” (except on the Packtpub site, where the more general “automated machine learning” was used due to smaller total publication volume). The collection was conducted from 20th of December 2021 till 3rd of March 2022. A funnel-type approach was used, narrowing down the results from the initial results list to a short list of relevant sources that were then used in the second phase of the research, building the model. The selection process is presented in Figure 2.

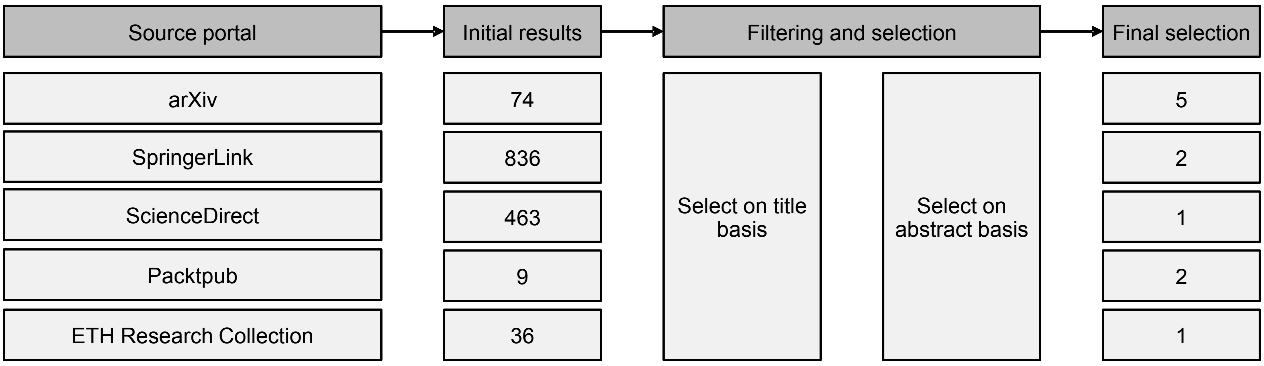


Figure 2: Literature sourcing and selection

The selection on title level was conducted by discarding all results that just presented or assessed the performance of underlying algorithms (e.g., “Performance review…”, “A new algorithm…”), which was most of the initial results. The remaining potential sources were screened on an abstract level to ensure that they were directly or indirectly describing features or functions of Auto-ML solutions. In this step, a lot of sources were discarded that used the term “feature” or “function” in a mathematical sense (e. g. “Feature space” or similar expressions). In total, eleven publications were selected to build the hierarchical function model. They are listed in Table 1. After deciding on the final selection, all mentions of functions or features of Auto-ML tools were extracted from the items in the table. The total number of mentions found in the respective publications are listed in the table as well (column “Number of functions”). All sources together yielded 275 functions and features (at this point including duplicates).

Table 1: Final selection of sources for the descriptive model

Source	Type	Sourced from platform	Number of functions
Truong et al. [10]	Journal paper	arXiv	31
Li et al. [17]	Journal paper	ETHZ	16
Humm, Zender [18]	Journal paper	Springerlink	5
Lee et al. [19]	Journal paper	Sciencedirect	8
Das [20]	Monograph (book)	Packtpub	42
He et al. [21]	Journal paper	arXiv	24
Hutter et al. [22]	Compilation (book)	Springerlink	22
Masood, Sherif [23]	Monograph (book)	Packtpub	42
Elshawi et al. [24]	Journal paper	arXiv	32
Zöller, Huber [25]	Journal paper	arXiv	41
Yao et al. [26]	Journal paper	arXiv	12

2.2 Building the hierarchical model

The model was built using the ARIS-toolkit by SCHEER [27], specifically its “function perspective”, which provides an interdisciplinary framework for modelling the functions of information systems and business processes for decision makers in IT as well as in management. To achieve the hierarchical structure of the model, which will help satisfy the individual information depth requested by stakeholders in potential integration projects, the functions are classified into four levels: function bundles, functions, partial functions and at the most granular level, elementary functions [28]. The syntax used in modeling the functions was "verb" + "noun" [28]. The first step of model building was removing all duplicate mentions of functions between the sources. Then, a preliminary classification of the functions was conducted, dividing them across the four levels according to SCHEER. Thirdly, the root node of the hierarchy was defined (the "function bundle" at the topmost level) and divided into functions, following the top-down approach recommended by SCHEER [27]. The structuring criterion used was the value creation process [27]. Following this, a bottom-up approach was used to connect the more granular functions to the above, aggregated layers. This combined approach has the benefit of a validation of the model's internal consistency. The final model comprises 149 functions divided across four levels.

3. Results

The 149 functions will be presented in nine parts, following the structure of the second level (the "functions"). The root node of the hierarchical model (the first level, called the "function bundle") is not visualized in the following figures, as it would simply repeat. This root node includes all subordinate functions and was named "Automate machine learning process" to provide a high amount of generality for the subordinate functions.

The first cluster starts at the data collection step. The main differentiating partial and elementary functions of Auto-ML solutions appear to be the different data types they can handle. Some are even fit to process unstructured data like images and videos (see Figure 3).

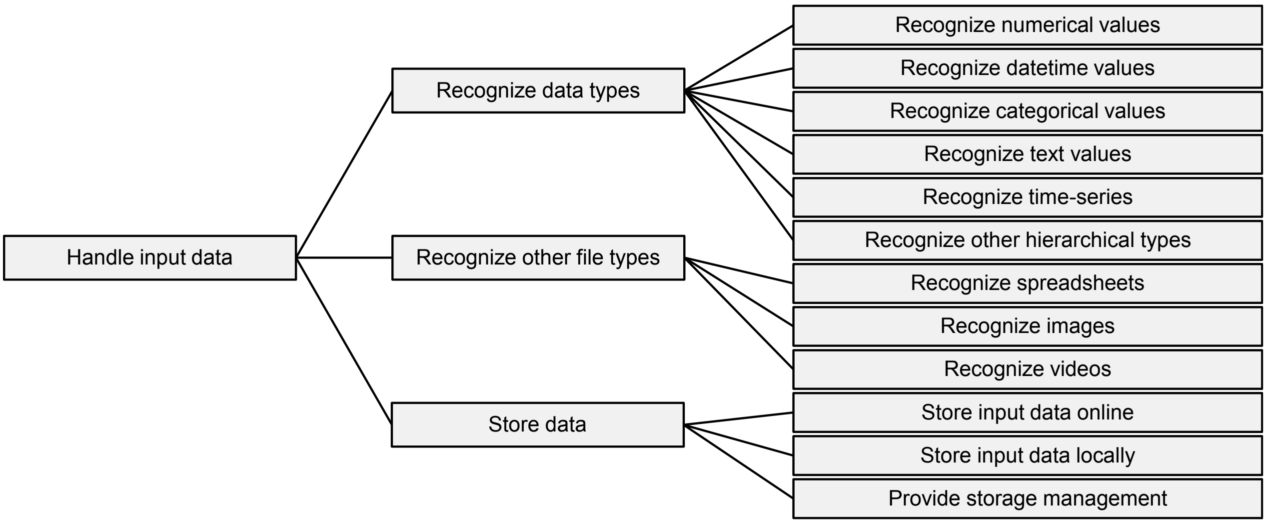


Figure 3: Handle input data

The second group comprises the functions concerned with the pipeline structure (see Figure 4). Three different approaches were found. The first (and most complex) function is only offered by a few tools. These tools can automatically create an entire machine learning pipeline and are not bound by a rigid or even sequential structure. The more common approach is to specify a fixed pipeline structure, in which the solution searches for the optimal pre-processing / model combination for the given problem. Lastly, some tools do not automate any tasks on the pipeline level but guide the user through the respective steps and recommend options that can be chosen.

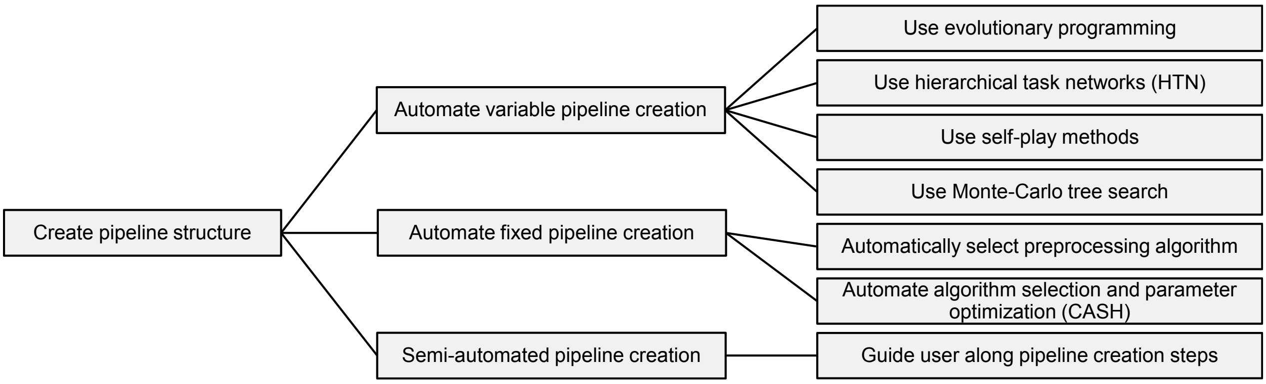


Figure 4: Create pipeline structure

The third cluster consists of partial functions that help prepare the stored data for further processing by cleaning unwanted anomalies, transforming and substituting data types and imputing missing values. This task is optionally simplified by some systems by visualizing the data in terms of different properties (see Figure 5).

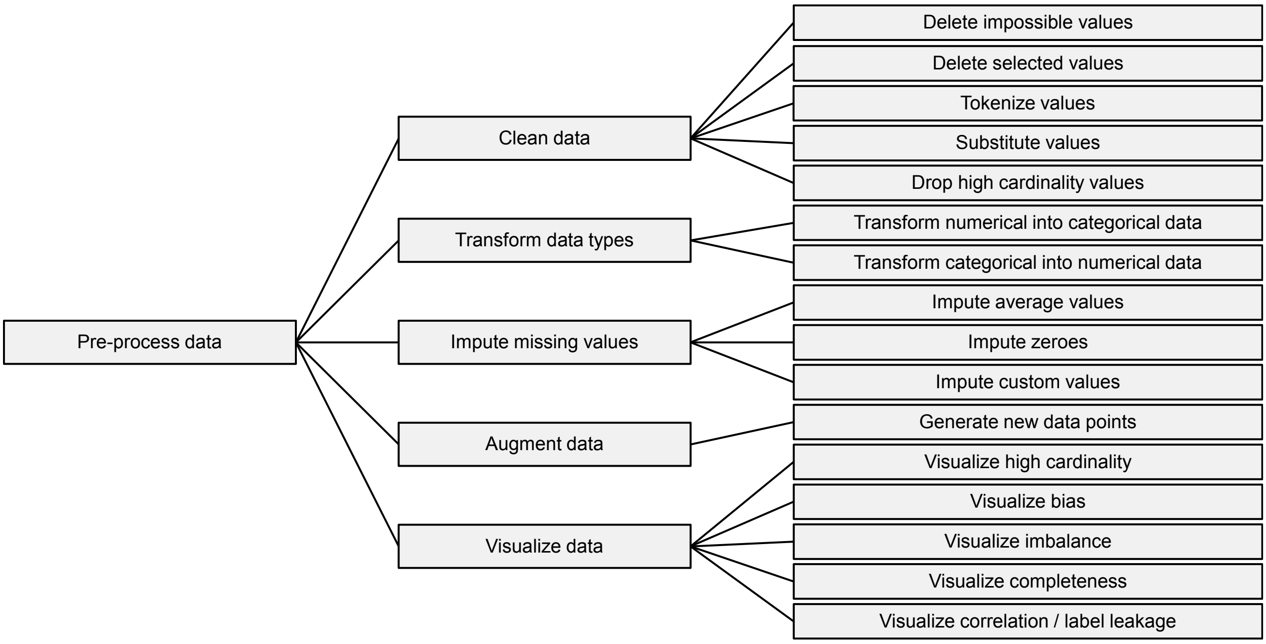


Figure 5: Pre-process data

The fourth cluster is occupied with feature engineering, a task that is particularly important for machine learning and normally requires experience in the field of data science. Well-selected features make the subsequent machine learning process more precise and improve the model’s performance. Here, the Auto-ML solutions support users by providing functions for rescaling and grouping and aiding in a pre-selection of possible features. Some also feature more advanced techniques like feature extraction, generation, or dimensionality reduction, which could help unexperienced users in training valid models on sparse, noisy, or highly dimensional datasets (see Figure 6).

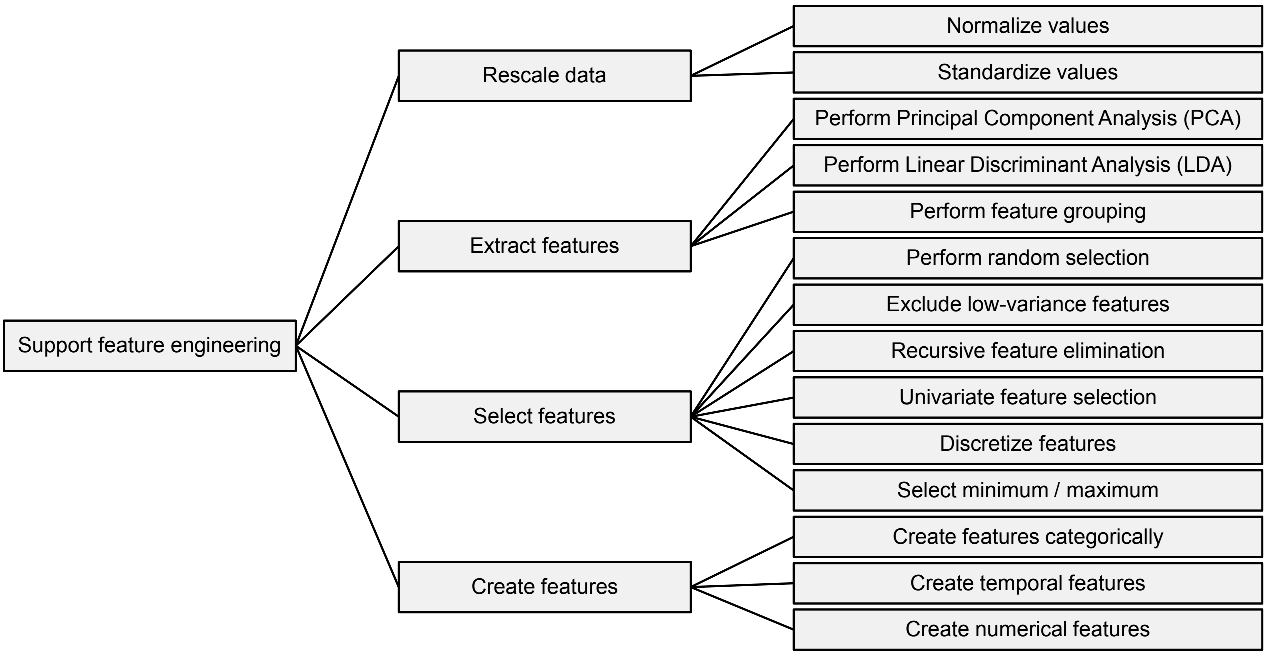


Figure 6: Support feature engineering

Approaching the actual modelling step, the research showed that different tools offer support for different kinds of machine learning problems. While the most popular problem type for Auto-ML appears to be supervised problems, some solutions offer a wider range of learning and modelling types (see Figure 7).

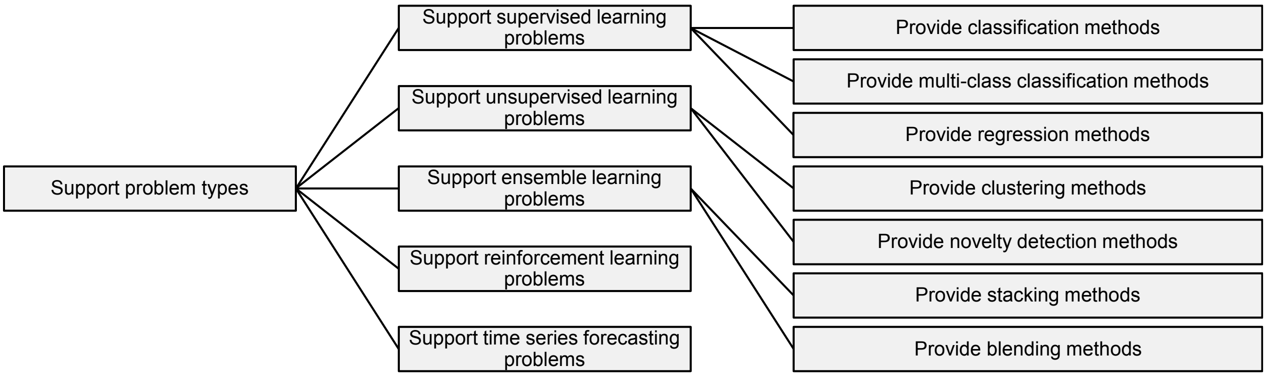


Figure 7: Support problem types

The sixth cluster is the biggest, as it represents the core back-end function of Auto-ML-Software. Here, the models are selected and their hyperparameters (i.e., the structural characteristics) are calculated based on the dataset (see Figure 8). A variety of techniques is employed by the different solution providers. Next to the variety of different model types, solutions also differ in their automation approach: Some choose combined algorithm selection and hyperparameter optimization (CASH) while others opt for letting the user choose a model and then optimizing the hyperparameters separately (conventional hyperparameter optimization or HPO). The generation of neural networks (NAS, short for neural architecture search) is closely related to HPO but can employ different search algorithms and thus is listed separately. Furthermore, some solutions provide quality-of-life-features (such as early stopping) to improve usability of the solution.

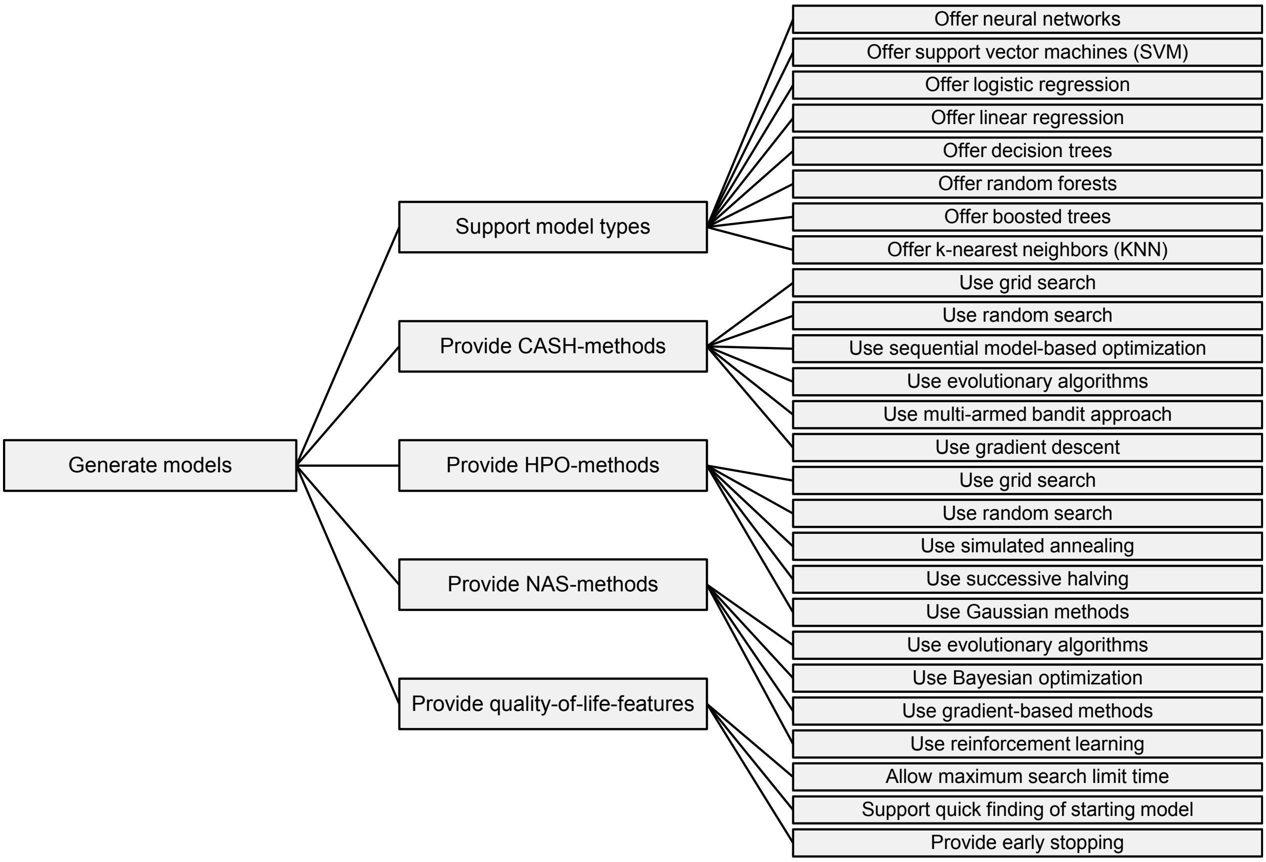


Figure 8: Generate models

The seventh cluster lists functions with which the user can review the results and choose a model to go forward with. The functions of Auto-ML software in this step revolve around helping the user make an informed decision about what is the optimal model for their use case. This can be done by ranking the models according to different criteria and visualizing their differences (see Figure 9).

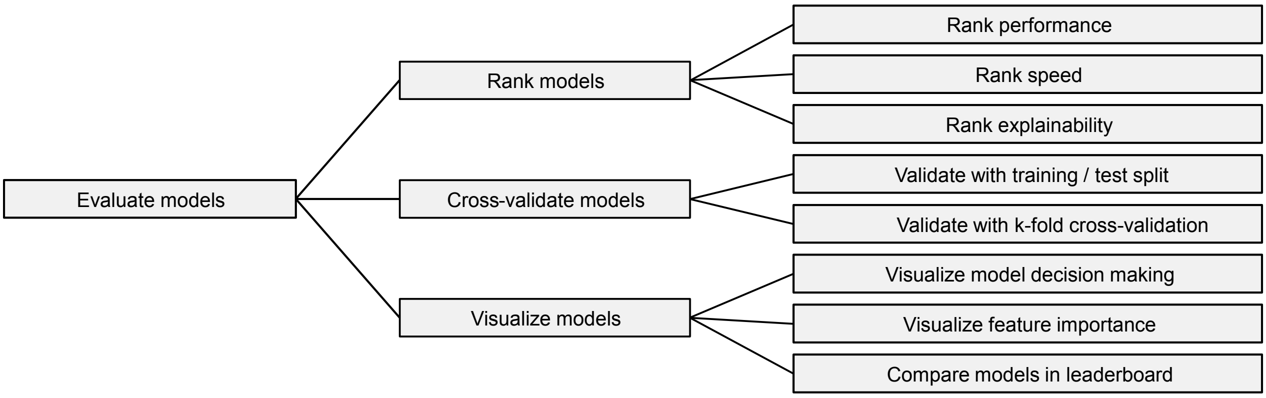


Figure 9: Evaluate models

After a model has been chosen and validated, some solutions provide support for end user in employing the models productively. Normally, this would demand some expertise in software engineering. However, some solutions even include their own cloud-based service that lets users host models online, significantly reducing the effort of providing the model as a service (see Figure 10).

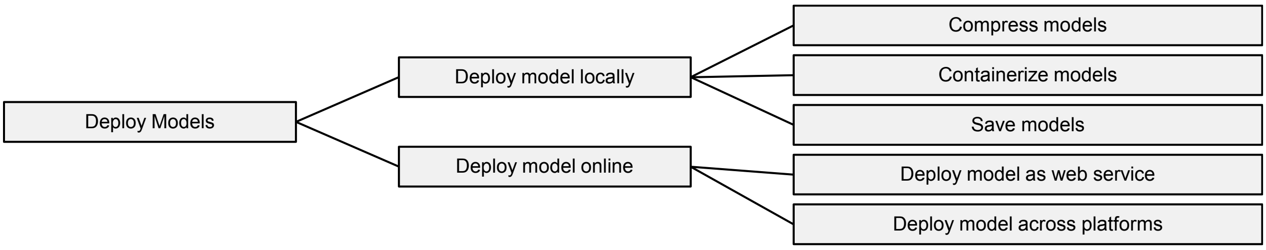


Figure 10: Deploy models

To offer support for users in the entire lifecycle of a machine learning model, some solutions even provide features for machine learning operations (ML-Ops), ensuring the correct and safe use of the model predictions in the field. Features include guardrails to minimize unexpected behaviour and providing model alerts to boost safety when using the model to steer sensitive processes. Lastly, meta-learning functionalities help to create new models more efficiently by learning best-practices from previous modelling efforts (see Figure 11).

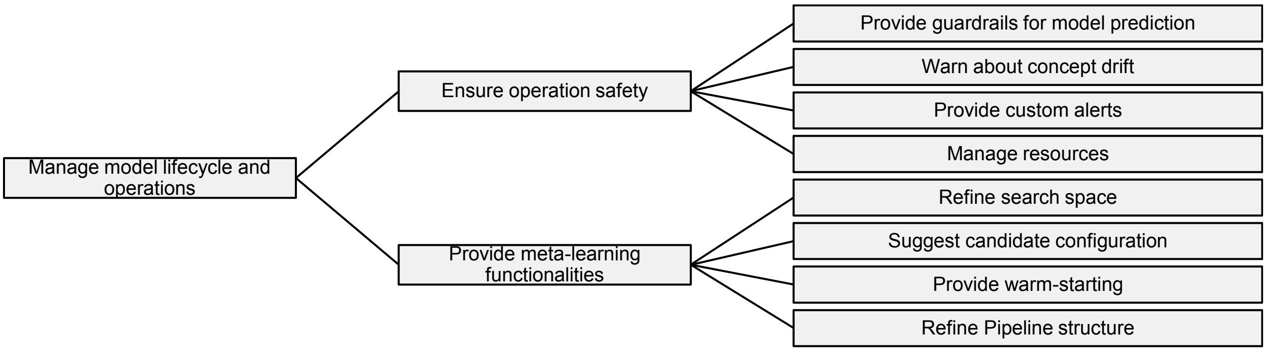


Figure 11: Optimize model operations

4. Discussion

The provision of automated machine learning software is an emerging, highly competitive market and thus the above results are subject to change. New functions may be implemented into established solutions, or a new competitor could enter the market and provide a disruptive set of features that changes the playing field. Thus, the descriptive model provided represents a snapshot of the feature sets available for a limited amount of time. However, care was taken to implement a modular model structure as well as a reproducible approach. This means that the model can easily be augmented and adapted in case of changes.

5. Summary and Outlook

The aim of the paper was to develop a new descriptive model of the functions of Auto-ML software solutions. A hierarchically structured model was chosen to give a dynamic depth of information. The information for building the model was sourced using the integrative literature review technique by TORRACO, while the model itself was built with the ARIS methodology described by SCHEER. Like mentioned in the introduction of the paper, the descriptive model of Auto-ML functions is only the first step to create a structured process, with which companies can select a fitting Auto-ML solution for their needs. To enhance practical usability, the user perspective must be considered as well. Companies have differing requirements and needs regarding machine learning, from performance to transparency and security. These research questions will be tackled in future publications.

Acknowledgements

This effort has been funded by the Leitmarkt.NRW program (EFRE-0801386) and the European Union in the European Regional Development Fund (EFRE) under the name “TechRad”. The authors wish to acknowledge the EFRE for their support.

References

[1] Reder, B., 2021. Studie Machine Learning 2021. https://www.lufthansa-industry-solutions.com/de-de/studien/idg-studie-machine-learning-2021. Accessed 2 September 2021.

[2] Chen, Y.-W., Song, Q., Hu, X., 2021. Techniques for Automated Machine Learning. SIGKDD Explor. Newsl. 22 (2), 35–50.

[3] Das, P., Perrone, V., Ivkin, N., Bansal, T., Karnin, Z., Shen, H., Shcherbatyi, I., Elor, Y., Wu, W., Zolic, A., Lienart, T., Tang, A., Ahmed, A., Faddoul, J.B., Jenatton, R., Winkelmolen, F., Gautier, P., Dirac, L., Perunicic, A., Miladinovic, M., Zappella, G., Archambeau, C., Seeger, M., Dutt, B., Rouesnel, L., 2020. Amazon SageMaker Autopilot: a white box AutoML solution at scale. http://arxiv.org/pdf/2012.08483v2.

[4] Statista Research Department, 2022. Number of AI/ML service offerings at hyperscale CSPs worldwide 2020-2021, by provider. Statista Research Department. https://www.statista.com/statistics/1268286/worldwide-ai-machine-learning-service-offerings-hyperscalers/. Accessed 7 March 2022.

[5] Kaul, A., Schieler, M., Hans, C., 2019. Künstliche Intelligenz im europäischen Mittelstand: Status quo, Perspektiven und was jetzt zu tun ist. https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/kaul/Universita%CC%88t_des_Saarlandes_Ku%CC%88nstliche_Intelligenz_im_europa%CC%88ischen_Mittelstand_2019-10_digital.pdf. Accessed 7 March 2022.

[6] Cetindamar, D., Phaal, R., Probert, D.R., 2016. Technology management as a profession and the challenges ahead. Journal of Engineering and Technology Management 41 (4), 1–13.

[7] Schuh, G., 2012. Innovationsmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg, 422 pp.

[8] Carlsson, K., Gualtieri, M., 2019. The Forrester New Wave™: Automation-Focused Machine Learning Solutions, Q2 2019: The Nine Providers That Matter Most And How They Stack Up, Cambridge, 19 pp. Accessed 17 February 2021.

[9] Dilmegani, C., 2022. AutoML Tech: Products of 2022 Compared: in-Depth Guide. https://research.aimultiple.com/automl-comparison/. Accessed 12 January 2022.

[10] Truong, A., Walters, A., Goodsitt, J., Hines, K., Bruss, C.B., Farivar, R., 2019 - 2019. Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools, in: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, USA. 04.11.2019 - 06.11.2019. IEEE, pp. 1471–1479.

[11] Torraco, R.J., 2005. Writing Integrative Literature Reviews: Guidelines and Examples. Human Resource Development Review 4 (3), 356–367.

[12] Cornell University, 2022. arXiv. https://arxiv.org/. Accessed 12 January 2022.

[13] Springer Verlag, 2022. SpringerLink. https://link.springer.com/. Accessed 12 January 2022.

[14] Elsevier B. V., 2022. Sciencedirect. https://www.sciencedirect.com/. Accessed 12 January 2022.

[15] Packt Verlag. Packtpub. https://www.packtpub.com/. Accessed 12 January 2022.

[16] ETH Zürich, 2022. ETH Research Collection. https://www.research-collection.ethz.ch/. Accessed 12 January 2022.

[17] Li, Y., Wang, Z., Ding, B., Zhang, C., 2021. AutoML: A Perspective where Industry Meets Academy, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. KDD '21: The

27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event Singapore. 14.08.2021-18.08.2021. ACM, New York, NY, USA, pp. 4048–4049.

[18] Humm, B.G., Zender, A., 2021. An Ontology-Based Concept for Meta AutoML, in: Maglogiannis, I., Macintyre, J., Iliadis, L. (Eds.), Artificial Intelligence Applications and Innovations, vol. 627. Springer International Publishing, Cham, pp. 117–128.

[19] Lee, K.M., Yoo, J., Kim, S.-W., Lee, J.-H., Hong, J., 2019. Autonomic machine learning platform. International Journal of Information Management 49, 491–501.

[20] Das, S., 2018. Hands-On Automated Machine Learning: A beginner's guide to building automated machine learning systems using AutoML and Python, 1st ed. Packt Publishing Limited, Birmingham, 282 pp.

[21] He, X., Zhao, K., Chu, X., 2021. AutoML: A Survey of the State-of-the-Art. Knowledge-Based Systems 212 (3), 106622.

[22] Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated Machine Learning: Methods, Systems, Challenges. Springer International Publishing; Imprint: Springer, Cham, 219).

[23] Masood, A., Sherif, A., 2021. Automated Machine Learning, 1st edition ed. Packt Publishing; Safari, Erscheinungsort nicht ermittelbar, Boston, MA, 312 pp.

[24] Elshawi, R., Maher, M., Sakr, S., 2019. Automated Machine Learning: State-of-The-Art and Open Challenges. http://arxiv.org/pdf/1906.02287v2.

[25] Zöller, M.-A., Huber, M.F., 2021. Benchmark and Survey of Automated Machine Learning Frameworks. Journal of Artificial Intelligence Research 70 (1), 409–474.

[26] Yao, Q., Wang, M., Chen, Y., Dai, W., Li, Y.-F., Tu, W.-W., Yang, Q., Yu, Y. Taking Human out of Learning Applications: A Survey on Automated Machine Learning.

[27] Scheer, A.-W., 1998. ARIS - Modellierungsmethoden, Metamodelle, Anwendungen, Dritte, völlig neubearbeitete und erweiterte Auflage ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 219 pp.

[28] Krcmar, H., 2015. Informationsmanagement, 6. überarb ed. Springer, Berlin [u. a.].

Biography

Günther Schuh (*1958) is head of the chair of Production Systems (WZL-PS) at RWTH Aachen University and member of the directorates of the Machine Tool Laboratory (WZL) at the RWTH, Fraunhofer Institute for Production Technology (IPT) and Director of the FIR at RWTH Aachen University.

Max-Ferdinand Stroh (*1991) is a researcher at FIR at RWTH Aachen University since 2017 in the department Information Management. He is head of the department Information Management. His scientific work is focused on the practical application of AI, smart products, and IT-OT-Integration.

Justus Aaron Benning (*1995) is a researcher at FIR at RWTH Aachen University in the department of Information Management since 2019. He is leading the group Information Logistics and is head of software development in his department. While his degree is in mechanical engineering and business administration, he spent a semester abroad at Korea University in Seoul to focus on the business applications of artificial intelligence.

Stefan Leachu (*1996) is a researcher at FIR at RWTH Aachen University in the department of Information Management since 2021. With his degree in computer science, he enjoys linking the technical issues of various technologies with the value-added potential that can be achieved in the business world.

Katharina Schmid (*1992) studied business economics at DHBW Ravensburg, specializing in logistics and controlling. After graduation, she worked as a project manager in supply chain projects for an international automotive company. To further improve her skills, she then studied computer science at RWTH Aachen University. She joined FIR at RWTH Aachen University in 2019 in the department of Information Management.

3rd Conference on Production Systems and Logistics

Characterization of Relationships in Data Ecosystems

Joshua Gelhaar[1], Felix Becker[1], Tobias Groß[1]
[1]*Fraunhofer ISST, Dortmund, Germany*

Abstract

The importance of data as a strategic resource for the development of innovation is steadily growing. Data-driven value creation increasingly requires cross-company collaboration between various actors with different roles in so-called data ecosystems. So far, however, the existing knowledge in the research field around data ecosystems is still relatively limited. In particular, the relationships and interdependencies between the different actors in a data ecosystem are not well understood yet. To address this research gap, we conduct a structured literature review and interview eleven experts from practice to identify characteristics of relationships between actors in data ecosystems. Among other things, the results show that both tangible characteristics, such as a clear exchange of values, and intangible characteristics, such as trust, are distinguishing features of the relationships between actors in data ecosystems. These study results can serve as a tool for both researchers and practitioners to better understand data ecosystems in general and the relationships and interactions that occur within them.

Keywords

Characterization; Relationships; Data Ecosystems; Data Sharing; Literature Review; Expert Interviews

1. Introduction

The ongoing digitization of the economy and society leads to large amounts of data. Advancing data analytics techniques are harnessing these data volumes and driving changes in existing businesses as well as the emergence of novel business opportunities [1]. At the same time, this development leads to internal company data increasingly being used externally and vice versa, and consequently to organisational boundaries being broken down [2]. These circumstances result in data-driven innovation and economic value creation being less and less created by individual organizations or in traditional value chains [3]. Instead, today's business world is becoming increasingly interconnected, combining various data sources from different organizations in cross-industry, sociotechnical networks – so-called data ecosystems [4]. Both researchers and practitioners argue that now and in the future, ecosystem participation is not a choice but a necessity for companies to take advantage of data sharing and remain competitive in the long term [5,4]. Despite that, many companies still refuse to share their data across companies and thus cannot use the potentials of data ecosystems for their benefits [6]. Especially in traditional sectors such as production and logistics, data is still comparatively rarely shared beyond company boundaries [7]. One rationale for this is seen in the lack of studies and the consequent absence of generally accepted theories and models for data ecosystems [8,4]. [9] conclude in their systematic review of the data ecosystem literature that more research is needed regarding relationships of actors and their characteristics in data ecosystems for theory development. This can be reasoned by the fact that the relationships between the actors create and span the ecosystem and ultimately determine how it functions [10]. In addition, understanding the types of relationships and interactions between the actors is an important step in building and developing a data ecosystem [11]. A better understanding of the actor's roles

DOI: https://doi.org/10.15488/12159

and their interrelationships can help to describe data ecosystems more precisely and formally and can serve as a basis for the development of a meta-model or an ontology, for example [12]. To the best of our knowledge, there is no scientific publication yet that deals with the detailed description of the characteristics of relationships in data ecosystems and addresses the research gap mentioned above. To contribute to a deeper understanding of the evolving research field around data ecosystems, we want to address this research gap and answer the following research question in this paper:

Research Question: *How can the relationships of actors in data ecosystems be described and characterized?*

To answer this research question, we develop an overview and description of the characteristics of relationships in data ecosystems. The results are grounded both in the scientific literature, through the conduct of a structured literature review, and in practice, with the help of expert interviews.

The remainder of this paper is structured as follows: After the introduction, we outline the theoretical background of data ecosystems, their roles and relationships, and draw a distinction to existing preliminary work. In Section 3, we outline our research approach and data collection processes. The characteristics of relationships in data ecosystems that we identified are described in section 4. The paper concludes with a discussion of the results and implications for research and practice.

2. Research background

2.1 Data ecosystems

The term ecosystem was coined by the biologist Arthur Tansley [13] who proposed a concept to describe the interactions between organisms of different species and their environment as an integrated system [14,15]. Based on this definition, various strands of research have since emerged in which the principles of the biological ecosystem concept are applied to other domains. One of the most popular concepts is the one of business ecosystems which were popularized by James Moore [16]. He uses this concept to describe interacting organizations as an "*economic community*" that aims to produce innovative products and services for customers, which are also part of the business ecosystem [16]. The increasing penetration of digital technologies within the business world has led to the analogy of digital ecosystems, which is seen as a "*digital version*" of business ecosystems [2]. The focus of this study lies on data ecosystems, which can be seen as a special type of digital ecosystems [17]. Following other ecosystem concepts, data ecosystems consist of diverse interactions between multiple actors that contribute to the creation and manipulation of a resource – which in this case are data – through joint activities [4]. On that basis, we see the focus of data ecosystems in the cross-actor generation, processing, sharing, and use of data with the goal to create added value for all actors involved [18,9].

Data ecosystems have certain characteristics that distinguish them from other forms of inter-organizational cooperation, such as traditional value chains or networks [19,15]. One characteristic is the lack of clear ecosystem boundaries, which can lead to varying degrees of interdependencies and relationships among the participating actors and ultimately to a heterogeneous and changing set of members [10]. Another characteristic is referred to as "*co-evolution*" [10]. It describes the condition that the development of one actor can positively affect the development of the other actors, resulting in benefits for all involved [15]. This is also because the ecosystem actors can have cooperative and competitive relationships at the same time – also known as coopetition [15,20].

2.2 Roles and relationships in data ecosystems

Based on the definition above, a data ecosystem consists of multiple actors. An actor is an autonomous entity, such as a company, an institution, or an individual person [4]. Depending on various factors, e.g., the motivation and capabilities of the actors, they perform different functions in the data ecosystem. A function or activity performed separately in this way is called a role within the ecosystem [18]. An individual actor may in turn perform one or more of these roles in a data ecosystem [21,12]. Exactly which roles can exist in a data ecosystem and which roles are essential is still debated in the literature [9]. However, there is a general agreement that there need to be at least three roles in a data ecosystem [22,23]. First, this is the role of the data provider who is responsible for the generation and collection of data [18,9]. Second, it requires a role that analyses and interprets data which is called analysis service provider [24]. Last, the information gained through data analysis is used by the data user role to generate value from it [25]. In addition to these three roles, the role of the so-called “keystone” actor is also frequently mentioned in the literature [9]. In some data ecosystems, this role may be responsible for providing most of the data as well as promoting the ecosystem, and thus may be instrumental in the ecosystem's growth and success [11,26]. Nevertheless, there are also data ecosystems that have a rather decentralized, distributed organizational form and thus operate without a central actor [8,27]. Instead, these ecosystems are held together by their common goal of shared value creation [11,9]. Since there is already some basic understanding about the roles in data ecosystems and several papers already exist on this topic (see e.g. [4], [9], or [18]), we do not focus on the detailed description of roles in data ecosystems in this paper. Instead, we concentrate on describing and characterizing the relationships between the data ecosystem actors. The reason for this is that the individual actors in a data ecosystem do not generate added value on their own. Rather, the added value arises through the interactions and relationships among each other, such as the exchange and sharing of data [11,12]. Building on [4], we see a relationship in a data ecosystem as an interaction between two data ecosystem actors, which is influenced by their roles and characterized by certain attributes. The detailed description of these characteristics is the goal of this paper.

In the scientific literature on data ecosystems, there is little prior work that has explicitly addressed the relationships and interactions between the different actors of a data ecosystem. Noteworthy in this context, however, are the works of [12], [8], and [21]. Based on a literature review, [12] developed a meta-model describing the basic concepts of data ecosystems and their relationships to each other. However, the authors only identify the essential elements such as actors, roles, relationships, and resources of data ecosystems and do not go into further detail or elaborate on characteristics of the individual elements such as the relationships. The study by [8], who developed a taxonomy for data ecosystems, has a similar focus. The taxonomy contains the essential characteristics and dimensions of data ecosystems but does not include specific characteristics of actor relationships. [21], conversely, examine in more detail the relationships and interactions in what they call data exchange ecosystems. However, the authors base their analysis on graph theory, which leads them to describe only structural features of relationships in data ecosystems, such as the number of relationships. We believe, by contrast, that this does not take into account all the characteristics of relationships, such as trust as an informal characteristic [27].

3. Research approach

To answer the research question formulated above, we aim to describe the characteristics of relationships in data ecosystems in a structured and concise way. For this, we conducted a structured literature review that is based on standardized and accepted guidelines in our research field (see section 3.1). To incorporate insights from practical reality into the research findings of this paper, we also conducted a series of eleven expert interviews (see section 3.2). In this way, we were able, on the one hand, to define and describe an initial number of characteristics through the literature analysis. On the other hand, we were able to validate and specify the characteristics identified in the expert interviews on the basis of the literature.

3.1 Literature review

We conducted a structured literature review following the approach described by [28] and the guidelines by [29]. Based on the research question formulated above, we chose Scopus as our scientific literature database because it contains more than 25,100 titles from more than 5,000 international publishers and thus promises very good results for our field of interest as it indexes the most relevant journals and conference proceedings [30]. We used the search string "data ecosystem" OR "data-driven ecosystem" OR "data-based ecosystem" as these terms are used synonymously by some authors [8]. The results were limited to English-language literature and peer-reviewed only. As a result, we received 457 as the initial set of publications. Within this first set, we reviewed the titles, abstracts, and keywords of the articles to check whether a hit fit the research scope. If the content of an article remained unclear, we examined the whole paper. We eliminated publications that did not fit our research scope or had no relevance with data ecosystems and the relationships between their roles. For example, we excluded papers dealing with the so-called Big Data Ecosystem (see e.g. [31]) as these often describe the software ecosystem around Hadoop and therefore have a different focus than the data ecosystems we intend to study. Eventually, this resulted in 64 relevant papers. In a second iteration, we complemented the literature set with a forward- and backward-search, as suggested by [28], to identify additional relevant publications. This led to the consideration of 12 additional useful articles. Consequently, a total of 76 articles were considered for the literature review. Figure 1 provides an overview of the literature search process. We analysed the publications thoroughly to extract scientific insights about the key features and characteristics of relationships between actors in data ecosystems from the body of literature.

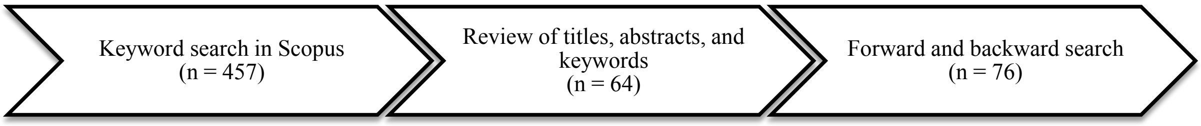


Figure 1: The structured literature review process

3.2 Expert interviews

To gain insights into the real-world environment, i.e., the relationships between actors within data ecosystems, we conducted a series of eleven interviews with different experts. Before conducting the interviews, we created an interview guide to ensure that all conversations covered a similar range of topics, characterizing the interviews as semi-structured [32]. All interviews were conducted remotely via Internet communication tools, recorded, transcribed, and anonymized. The analysis of the transcripts followed a qualitative context analysis [33]. To support our findings in this work, we follow the recommendation of [34] and cite power and proof quotes in the presentation of our results, i.e. quotes that strongly underline the point we are trying to make.

The selected interviewees come from various industries and functions such as management, project and operations managers, and consultants (see Table 1). The selection process took care to interview representatives from organizations that actively take one or more roles in a data ecosystem as well as representatives from organizations that advise other companies on how to engage in data ecosystems. This ensured that both an inside view and an outside view of the characteristics of relationships between roles in data ecosystems were considered. The interviewees were asked questions about the general understanding of data ecosystems, the relationships and interdependencies between the participating ecosystem roles, as well as the potentials and challenges that can arise from these relationships. These questions were particularly concerned with understanding how and why organizations establish relationships with other actors in data ecosystems.

Table 1: Overview of the interviewed experts

#	Organization / Industry	Position	Duration [min:ss]
E1	Industry association	IT Specialist	42:15
E2	Conglomerate	Senior Director	47:39
E3	Automotive	IT Manager	32:24
E4	Consulting	Consultant	27:21
E5	Consulting	Consultant	26:53
E6	IT Consulting	Senior Consultant	46:25
E7	Industry association	Innovation Advisor	39:13
E8	Conglomerate	Senior Principal Engineer	40:21
E9	Manufacturing	Principal Consultant	36:12
E10	Academy of sciences	Scientific Officer	33:37
E11	Manufacturing	Board Member	44:28

4. Results

In this section, we present the results of our methodological approach. Overall, we were able to identify the six characteristics Value Stream, Co-Creation, Interdependence, Trust, Intention, and Relevance. These characteristics are explained in detail in the following:

Value Stream – Every relationship in a data ecosystem is based on a value stream. By definition, within data ecosystems, these value streams focus on the exchange and sharing of data. However, both the literature and the expert interviews made it clear that data providers do not necessarily have to receive data in return for their data offering (see e.g. [22]). Instead, data providers can be compensated for their data in various ways. Expert 10 put it as follows: "*And in this data ecosystem, the people interact in some way and then have some kind of interaction with each other. Most of the time, the data goes from one person to another, who perhaps pays money for it or can offer a service for it…*". However, the simplest form of compensation is money, as money is easy to handle and applicable in many situations [35]. At the same time paying money implies the challenge of valuing data in monetary terms [36]. This is one of the reasons why data providers often receive a data-driven service in return for their data instead of money [18]. An example that was mentioned by some experts was the use case in which a machine operator makes his data available to a service provider and receives a predictive maintenance service in return. Consequently, the value stream between two actors in a data ecosystem consists of the exchange of data, money, or services. It is usually easy to determine which value is exchanged between the roles, but it is often difficult to say whether they have the same (monetary) value.

Co-Creation – The second identified characteristic of relationships in data ecosystems is the strong degree of collaboration between the actors to jointly create value. In the literature, this is also referred to as value co-creation [18,3]. This is due to the fact that in data ecosystems, various actors converge, who can offer different data, and because they may have varying data analysis capabilities [11,9]. Only through close collaboration between these ecosystem actors value can ultimately be created for all involved stakeholders [37]. In the interviews with the industry experts, it became clear that in all data ecosystems, one actor cannot take on all the necessary roles and tasks, and therefore the collaboration of different actors is inevitable for the realization of most use cases. "*In a data ecosystem companies are engaged in value-adding activities based on data and jointly develop new service-providing products or services*" said expert 5. In turn, expert 1 emphasizes that this collaboration is particularly relevant for small and medium-sized enterprises as their size makes them dependent on other companies to be able to exploit the potential arising from data ecosystems. The degree and quality of collaboration between two actors are difficult to quantify. However, an indication of a high degree of collaboration could be the frequency of communication, i.e. the exchange of data and information between the actors [21].

Interdependence – Related to the characteristic co-creation is the characteristic interdependence. It describes the degree to which different actors in the ecosystem must interact and share data to perform their roles and tasks [38]. Expert 4 has expressed this as follows: "*Characteristic of a data ecosystem is that the actors are dependent on each other in some way. For some companies, this dependence on others may be stronger than for others.*" Analogous to other ecosystem types, this interdependence can vary [8]. Close and highly dependent relationships are called tightly coupled [39]. Loosely coupled refers to relationships that are less formal and close, suggesting a higher degree of openness in the ecosystem [15]. For example, [21] showed in their study that the same roles in a data ecosystem usually have little or no relationship to each other and are therefore relatively independent. The authors explain their findings with the fact that the same roles in a data ecosystem can be in competition with each other and therefore do not want to exchange data with each other, for example. However, a dependency does not have to be bilateral. [26] point out, for example, that in some cases service providers can be very dependent on certain data providers, and therefore, there may be more of a one-sided dependency.

Trust – It became apparent in both the literature review and the expert interviews that trust is a central characteristic of any relationship in a data ecosystem. [40] state, for example, "*Trust is the fundamental component of all relationships in a data sharing ecosystem*". In data ecosystems, this usually involves trust in how other actors handle the data of the data providers and whether they adhere to the agreed terms of use [18]. Expert 2 highlighted the following in this regard: "*But that means you have to trust your supplier not to, let's say, play fast and loose with your data and pass it on to some other competitor and say to them, "look how they do it in their factory".*" However, it is difficult to say how trust takes shape in a relationship since trust is an informal factor that can be difficult to measure [17,27]. Nevertheless, theory and practice agree that relationships in data ecosystems are often built on pre-existing business relationships [11]. Furthermore, in a data ecosystem, various measures such as technological design decisions and formal agreements can be taken to increase trust in the data exchange and thus in the relationships between the actors [27,6]. Examples of this are the use of technologies that allow data usage conditions to be defined and enforced, such as the International Data Spaces [11], or the use of blockchain technologies that can replace trust in an intermediary [41].

Intention – Every actor in an ecosystem has an incentive and motivation for their participation [15,21]. A relationship in an ecosystem can only develop if the motivation of both actors matches together [42,4]. However, this is not trivial, as the motivation of an actor is not always clear for the other actors, can be contradictory in some areas due to competition, or can also change over time [22,11]. A relationship in a data ecosystem is therefore characterized by a constant tension between the motivations of the actors. Several experts emphasized in the interviews that a relationship between two actors can only emerge and develop if the added value of the relationship, i.e. how they can benefit from the collaboration, is clear to both sides. Expert 11 states in this regard in the interview: "*If in the end not everyone benefits, companies would not be interested. This is one of the basic principles of the industry: "I don't do anything I can't profit from".*"

Relevance – Looking at a data ecosystem as an interconnected system, it can be noted that some relationships are more important to the emergence and functioning of the ecosystem than others. This can be explained by the fact that some actors bring resources such as data or services into the ecosystem that cannot be provided at all or only with difficulty by other actors [18]. The relationships that these actors have with other actors, especially if these are relationships between important actors, may therefore be more relevant to the functioning of the ecosystem than others [11,21]. Typically, these are the relationships between the data providers, who offer much of the relevant data, and the analysis service providers, who offer the core service of the data ecosystem [18,43]. Conversely, some relationships provide more supportive services to the ecosystem and are easier to replace. Examples of support services are data quality evaluation or matchmaking services [43]. A frequently mentioned support service is also the so-called infrastructure provider [18,17]. However, it became clear in the interviews that this service is difficult to categorize uniformly in terms of its relevance, since the infrastructure can have a high or rather low relevance depending on the protection requirements of the shared data. For example, expert 2 said "*The technology used for data sharing can be important in some data ecosystems, but not in others. For example, when it comes to possible areas of competition, i.e. company X may not be able to run its products over an architecture from company Y.*" For this reason, among others, we believe that the relevance of a relationship in a data ecosystem is highly context-dependent and can only be determined through qualitative analysis.

5. Conclusion

This study deals with the description and characterization of relationships between actors in data ecosystems. A precise understanding of the relationships is important as the interactions between the actors span the data ecosystem and it is only through collaboration that the added value of the ecosystem is created. Against this background, this study developed characteristics which describe relationships in data ecosystems in a structured and concise way.

Several implications for research and practice arise from the results of this work. From a **scientific perspective**, our results contribute to the still relatively young data ecosystem literature. While a few previous studies have dealt with the description and characterization of data ecosystems in general, the results of this study go deeper by describing the relationships between actors in more detail and more specifically. Consequently, the results, derived from the scientific knowledge base and expert interviews, can help to expand the existing body of knowledge and specify the common understandings and definitions of data ecosystems. Ultimately, this can be a further step towards developing fundamental and comprehensive theories of data ecosystems that do not currently exist in the scientific literature [8,9].

From a **managerial perspective**, the study results can be used by practitioners as a starting point to better understand the relationships of the data ecosystem in which the organization already participates. Based on this, relationships could be better managed and actively shaped. Ultimately, the findings of the study can help organizations develop relationships with other actors to build new data ecosystems to realize the potential of cross-organizational data sharing in data ecosystems.

However, our study is naturally subject to certain **limitations** which must be considered when interpreting the results. First, the steady progress of digitization and the still young age of the data ecosystem literature leads to evolving concepts and definitions around data ecosystems. Linked to this lack of commonly accepted theories is the challenge of distinguishing data ecosystems from related ecosystem concepts such as digital and platform ecosystems and associated forms of collaboration such as alliances and networks. Furthermore, data collection and analysis are also subject to certain limitations. On the one hand, additional articles could be found as data sources by using further scientific databases besides Scopus. On the other hand, the statements from the expert interviews are limited in their generalizability as other experts might give different answers. Finally, the analysis of the literature, as well as the interviews is subject to some interpretation, which means that other researchers may identify different characteristics depending on their influences, preferences, and biases.

However, the aforementioned limitations also point to possibilities for **future research** paths. By combining existing role descriptions and the results of this work, an ontology for data ecosystems could be developed that represents the essential roles and their relationships to each other in a structured way. An ontology would create a shared understanding of the topic of data ecosystems and be a further step towards developing a comprehensive theory [9]. Furthermore, it would be interesting to investigate how some characteristics can be built or developed. For example, what can be done if the intentions of two actors do not match? This may require incentive mechanisms in the data ecosystem that motivate actors to actively participate in the ecosystem and share their data [17,42].

6. References

[1] Brynjolfsson, E., McAfee, A., 2017. The Business of Artificial Intelligence: What it can — and cannot — do for your organization. Harvard Business Review 95, 3–11.

[2] Lis, D., Otto, B., 2020. Data Governance in Data Ecosystems - Insights from Organizations, in: AMCIS 2020 Proceedings.

[3] Hein, A., Weking, J., Schreieck, M., Wiesche, M., Böhm, M., Krcmar, H., 2019. Value co-creation practices in business-to-business platform ecosystems. Electron Markets 29 (3), 503–518.

[4] Oliveira, M.I., Lóscio, B.F., 2018. What is a data ecosystem?, in: Proceedings of the 19th Annual International Conference on Digital Government Research, Delft, Netherlands.

[5] Capgemini, 2021. Data Sharing Masters: How smart organizations use data ecosystems to gain an unbeatable competitive edge, 56 pp. https://www.capgemini.com/wp-content/uploads/2021/07/Final-Web-Version-of-Report-Data-Ecosystems.pdf.

[6] Prieëlle, F. de, Reuver, M. de, Rezaei, J., 2020. The Role of Ecosystem Data Governance in Adoption of Data Platforms by Internet-of-Things Data Providers: Case of Dutch Horticulture Industry. IEEE Trans. Eng. Manage., 1–11.

[7] Priego, L.P., Osimo, D., Wareham, J., 2019. Data sharing practice in Big Data ecosystems. Esade Working Paper Nº 273, 38 pp.

[8] Gelhaar, J., Groß, T., Otto, B., 2021. A Taxonomy for Data Ecosystems, in: Proceedings of the 54th Hawaii International Conference on System Sciences.

[9] Oliveira, M.I., Barros Lima, G.d.F., Lóscio, B.F., 2019. Investigations into Data Ecosystems: a systematic mapping study. Survey Paper. Knowledge and Information Systems.

[10] Jacobides, M.G., Cennamo, C., Gawer, A., 2018. Towards a theory of ecosystems. Strat Mgmt J 39 (8), 2255–2276.

[11] Gelhaar, J., Otto, B., 2020. Challenges in the Emergence of Data Ecosystems, in: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS).

[12] Oliveira, M.I., Oliveira, L.E., Batista, M.G.R., Lóscio, B.F., 2018. Towards a Meta-model for Data Ecosystems, in: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age, Delft, Netherlands, pp. 1–10.

[13] Tansley, A.G., 1935. The Use and Abuse of Vegetational Concepts and Terms.”. Ecology 16 (3), 284–307.

[14] Chapin III, F.S., A. Matson, P., M. Vitousek, P., 2014. The ecosystem concept, in: Chapin III, F.S. (Ed.), Principles of Terrestrial Ecosystem Ecology. Springer, Berlin, Germany, pp. 3–22.

[15] Guggenberger, T.M., Möller, F., Haarhaus, T., Gür, I., Otto, B., 2020. Ecosystem Types in Information Systems, in: Proceedings of the 28th European Conference on Information Systems.

[16] Moore, J.F., 1993. Predators and Prey: A New Ecology of Competition. Harvard Business Review 71 (3), 75–86.

[17] Cappiello, C., Gal, A., Jarke, M., Rehof, J., 2019. Data Ecosystems: Sovereign Data Exchange among Organizations: Report from Dagstuhl Seminar 19391. Dagstuhl Reports 9 (9), 66–134.

[18] Azkan, C., Möller, F., Meisel, L., Otto, B., 2020. Service Dominant Logic Perspective on Data Ecosystems - A Case Study based Morphology, in: Proceedings of the 28th European Conference on Information Systems.

[19] Adner, R., 2017. Ecosystem as Structure. Journal of Management 43 (1), 39–58.

[20] Nalebuff, B.J., Brandenburger, A.M., 1997. Co-opetition: Competitive and cooperative business strategies for the digital economy. Strategy & Leadership 25 (6), 28–33.

[21] Hayashi, T., Ishimura, G., Ohsawa, Y., 2021. Structural Characteristics of Stakeholder Relationships and Value Chain Network in Data Exchange Ecosystem. IEEE Access 9, 52266–52276.

[22] Gelhaar, J., Gürpinar, T., Henke, M., Otto, B., 2021. Towards a Taxonomy of Incentive Mechanisms for Data Sharing in Data Ecosystems, in: PACIS 2021 Proceedings, Dubai, UAE.

[23] van den Homberg, M., Susha, I., 2018. Characterizing Data Ecosystems to Support Official Statistics with Open Mapping Data for Reporting on Sustainable Development Goals. IJGI 7 (12).

[24] Azkan, C., Iggena, L., Gür, I., Möller, F.O., Otto, B., 2020. A Taxonomy for Data-Driven Services in Manufacturing Industries, in: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS).

[25] Curry, E., 2016. The Big Data Value Chain: Definitions, Concepts, and Theoretical Approaches, in: Cavanillas, J.M., Curry, E., Wahlster, W. (Eds.), New Horizons for a Data-Driven Economy. Springer International Publishing, Cham, pp. 29–37.

[26] Heimstädt, M., Saunderson, F., Heath, T., 2014. Conceptualizing Open Data ecosystems: A timeline analysis of Open Data development in the UK. Discussion Papers 2014/12, Free University Berlin, School of Business & Economics.

[27] Lis, D., Otto, B., 2021. Towards a Taxonomy of Ecosystem Data Governance, in: Proceedings of the 54th Hawaii International Conference on System Sciences, pp. 6067–6076.

[28] Webster, J., Watson, R.T., 2002. Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26 (2), xiii–xxiii.

[29] vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process, in: Proceedings of the 17th European Conference on Information Systems, Verona, Italy.

[30] Elsevier, 2020. Content - How Scopus Works. https://www.elsevier.com/solutions/scopus/how-scopus-works/content.

[31] Sumbaly, R., Kreps, J., Shah, S., 2013. The "Big Data" Ecosystem at LinkedIn, in: Proceedings of the 2013 international conference on Management of data - SIGMOD '13, New York, USA, pp. 1125–1134.

[32] Merton, R.K., Kendall, P.L., 1946. The Focused Interview. American Journal of Sociology 51 (6), 541–557.

[33] Krippendorff, K., 2013. Content analysis: An introduction to its methodology, 3rd ed. Sage, Los Angeles, 441 pp.

[34] Pratt, M.G., 2008. Fitting Oval Pegs Into Round Holes. Organizational Research Methods 11 (3), 481–509.

[35] Badewitz, W., Kloker, S., Weinhardt, C., 2020. The Data Provision Game: Researching Revenue Sharing in Collaborative Data Networks, in: 22nd Conference on Business Informatics (CBI), Antwerp, Belgium, pp. 191–200.

[36] Spiekermann, M., Wenzel, S., Otto, B., 2018. A Conceptual Model of Benchmarking Data and its Implications for Data Mapping in the Data Economy, in: Multikonferenz Wirtschaftsinformatik (MKWI). Data driven X - Turning Data into Value, pp. 314–325.

[37] Wilson, B., Cong, C., 2021. Beyond the supply side: Use and impact of municipal open data in the U.S. Telematics and Informatics 58, 101526.

[38] Curry, E., Sheth, A., 2018. Next-Generation Smart Environments: From System of Systems to Data Ecosystems. IEEE Intell. Syst. 33 (3), 69–76.

[39] Rong, K., Lin, Y., Li, B., Burström, T., Butel, L., Yu, J., 2018. Business ecosystem research agenda: more dynamic, more embedded, and more internationalized. Asian Bus Manage 17 (3), 167–182.

[40] Abebe, R., Aruleba, K., Birhane, A., Kingsley, S., Obaido, G., Remy, S.L., Sadagopan, S., 2021. Narratives and Counternarratives on Data Sharing in Africa, in: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency FAccT, pp. 329–341.

[41] Bons, R.W., Versendaal, J., Zavolokina, L., Shi, W.L., 2020. Potential and limits of Blockchain technology for networked businesses. Electron Markets 30 (2), 189–194.

[42] Gelhaar, J., Both, J., Otto, B., 2021. Requirements For Incentive Mechanisms In Industrial Data Ecosystems, in: Proceedings of the Conference on Production Systems and Logistics: CPSL 2021, Online. 10.08.2021 – 11.08.2021.

[43] Immonen, A., Ovaska, E., Paaso, T., 2018. Towards certified open data in digital service ecosystems. Software Qual J 26 (4), 1257–1297.

Biography

Joshua Gelhaar, M.Sc. (*1991) studied industrial engineering at TU Dortmund University. Since 2019 he works as a scientist in the department "Data Business" at the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund, Germany. His main research focus is on data sharing, especially within industrial data ecosystems.

Felix Becker, M.Sc. studied industrial engineering at TU Dortmund University and recently completed his master's thesis in the field of "Interdependencies and potentials of data ecosystems" in cooperation with the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund. He now works for a manufacturer of logistics systems.

Tobias Groß, M.A. (*1992) studied business management at FH Dortmund University of Applied Sciences and Arts. Since 2021 he works as a scientist in the department "Data Business" at the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund, Germany. His main research focus is on data sharing, especially investigating incentives for companies to take part in industrial data ecosystems.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2022

3rd Conference on Production Systems and Logistics

Towards A Methodology For Economic Performance Increase Of Production Lines Using Reinforcement Learning

Günther Schuh, Andreas Gützlaff, Judith Fulterer, Jan Maetschke
Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University

Abstract

The increasing number of variants in product portfolios contributes to the challenge of efficient manufacturing on production lines due to the resulting small batch sizes and thus frequent product changes that lower the average overall plant effectiveness. Especially for companies that manufacture at high speed on production lines, such as in the Fast Moving Consumer Good (FMCG) industry, it is a central task of operational management to increase the performance of production lines. Due to the multitude of different adjustment levers at several interdependent machines, the identification of efficient actions and their combination into economic improvement trajectories is challenging. There is a variety of approaches to address this challenge, e.g. simulation-based heuristics. However, these approaches mostly focus on details instead of giving a holistic perspective of the possibilities to improve a production line or are limited in practical application.

In other areas of application, reinforcement learning has shown remarkable success in recent years. The principle feasibility of using reinforcement learning in this application context has been demonstrated as well. However, it became apparent that the integration of expert knowledge throughout the improvement process is necessary. For this reason this paper transforms five modules defined from an engineering point of view into the mathematical scheme of a markov decision problem, a default framework for reinforcement learning. This provides the foundation for applying reinforcement learning in combination with expert knowledge from an engineering perspective.

Keywords

Production Lines; Production Management; Reinforcement Learning; Discrete Event Simulation; Performance; Effectiveness; Economic Efficiency; Markov Decision Problem; OEE

1. Introduction and challenges in increasing the performance of production lines economically

The high product variance demanded by the market combined with steadily increasing cost pressure and price sensitivity are raising the demands on the management of production to achieve business success. The resulting small batch sizes and frequent product changes lead to a reduction in average overall equipment effectiveness (OEE) [1–3]. This applies particularly to production lines, which can be found for example in the fast moving consumer goods (FMCG) industry, are characterized by generally high production speed and low margins [1]. Furthermore, companies allocate products in production networks back to western countries due to a higher standard of digitalization [4]. This combination results in consolidation and hence in increased planned utilization of production lines. As a result, the demands on the productivity and stability of production lines are rising. For this reason, a focus on the topic of performance increase in industry and research is perceived. [5,6].

DOI: https://doi.org/10.15488/12138

OEE, the productivity and stability of production systems and ultimately the production costs depend strongly on the configuration of production lines, consisting of several machines, buffers, conveyors, etc.[7]. Improving these production lines is a complex problem and the complexity increases drastically with the number of involved aggregates. The buffer allocation sub-problem on its own is an NP-hard problem [8,9]. As these problems can not be solved analytically, Discrete Simulation-Based Optimization (DSBO) is widely used in the industry to improve the configuration of production lines [10,9]. However, Studies show that companies need support in conducting precisely this DSBO studies, interpreting their results and deriving feasible step-by-step actions from them [11,12].

Besides this complexity of such systems, which makes optimization per se demanding, the identification of effective adjustment levers is challenging because the restraining element of the system shifts dynamically, due to the mutual dependencies of the system's elements. Additionally, not only the system's output is difficult to describe, but also the input in terms of efforts made, which converts to costs. The OEE only represents the output, but does not consider the input to achieve this output. That's why the identification and prioritization of economic actions for improvement only makes sense by considering the overall system behaviour and costs, not only by focusing on the bottleneck-orientated OEE [5,13–15].

The combination of several small actions on different machines is expected to yield higher efficiency gains than a major improvement on a single machine [16], an isolated consideration of sub-problems is therefore of limited benefit [14,9,15]. For this reason, economic performance considerations must focus specifically on the combination of individual measures.

Looking for decision support in such complex but well-defined optimization problems, artificial intelligence (AI) methods, especially reinforcement learning (RL), receive increasing attention in the last years [17,18]. The motivation for applying RL is that the RL agent learns to react efficiently to the dynamics of the environment, without any prior knowledge of the system dynamics [19].

This paper presents a method for increasing the performance of production lines in an economic and practical way using RL. The focus is not the demonstration of technical feasibility, which [13] already showed, but the integration of RL into a holistic improvement methodology. The aim is to discover trajectories of sequential improvements, which could be interpreted by engineers and implemented successively, but not to find optimal parameter settings. The method intends to provide practical decision support in individual cases without losing the character of a generalistic method.

The work is structured as follows. Section 2 discusses existing approaches in terms of meeting these challenges. *Section 3* shows the opportunities of combining RL and DSBO in this application context. *Section 4* then presents an approach combining these two technologies taking domain-specific knowledge into account. *Section 5* provides a summary and outlines further planned research activities.

2. State of the art

The following literature search is based on the procedure according to BORREGO ET AL [20], was conducted to identify an overview of previous approaches. First, a search string including synonyms is defined, see **Figure 1**. To ensure a broader search, no narrowing word related to *economic* increase is included in the string. This search string is used in the following search engine*s ScienceDirect, Web of Science, IEEE Explore, Scopus, Google Scholar* and returned 765 results. Removing duplicates resulted in 431 unique papers. Based on the title the number of relevant papers is reduced to 151. In the next step, figures and abstracts of the remaining papers are reviewed, resulting in a final 51 papers to be considered.

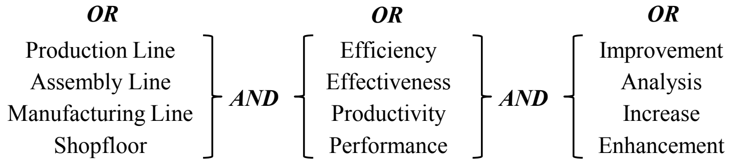


Figure 1: Search words used resulting in 64 search strings by combination

The majority (34 out of 51, 67%) of these papers do not present a methodology, but a case study on a unique application. Thus, these approaches can not be generalized and applied to other similar problems. Based on the other approaches presenting a method, additional papers of relevance are identified using the snowballing method. The combined results are discussed below.

The approaches can be roughly divided into mathematical and simulation-based approaches [21]. Even though mathematical approaches cannot cover the complexity of real use cases [22,10,9], a variety of specific analytical models for sub-problems exist [23,24]. Especially the optimization of buffer allocation has received much attention from researchers [25], such as in [26,8].

[14,9,15] argue that optimization is only possible by considering the entire system and not by focusing on improvement actions in such a specific way as the analytical models do, due to the complex dependencies of aggregates of production lines. At the same time, after a certain point, optimizing cycle times of individual aggregates is more economical than further improving the availability of all machines [14].

[1,27–29] explicitly consider fill-and-pack lines in the FMCG industry. However, they do not present an optimization approach, but rather simulation case studies as mentioned above. They underline the potential of optimizing such lines and show the need for a combined consideration of improvement costs and increased performance nevertheless and thus underline the motivation above.

[21,30,31,15] show that without considering the overall system, prioritizing improvement activities such as maintenance activities is not advisable and that this is not adequately addressed in the literature. None of the approaches listed systematically considers improvement trajectories, i.e. a sequence of independently realizable actions to improve a production system. Rather, they focuses on finding an (near-) optimal overall solution rather than looking at the path to get there, i.e. the improvement trajectories.

[12] gives an overview of DSBO approaches in manufacturing in general and shows that machine learning approaches for optimizing production systems are getting more and more attention in research. [10] sees the need for further research combining statistical learning in combination with DSBO. [17] predicts a vast increase in the importance of automated decisions based on AI in production management.

Due to the fact, that the improvement of production lines has been the subject of research for decades, the discussion above can only be a short summary. For more detailed references, the reader is advised to refer to [21,25,12,9].

In summary, there is a lack of approaches that provide practical support for improving the performance of production lines while considering the inherent complexity. As described in *Section 1*, reinforcement learning offers new methods to meet this challenge. The following paragraph discusses these opportunities.

3. Chances of combining reinforcement learning with simulation for the improvement of production lines

Discrete-event simulations (DES) are suitable for the evaluation of complex, stochastic systems, where a closed-form mathematical model is hard to find. Simulation is not an optimization technique itself and needs to be combined with optimization methods to improve problems of the real world [22]. It is advisable to statistically extract information from existing simulation runs to guide the parameter search and thus to closely integrate the optimization with simulation [32,10] Thus, optimization methods may need to be adapted to the specific problems [33,32].

For this reason, more and more approaches use AI for optimization in combination with simulation [12]. What makes RL a promising solution candidate is that it does not require holistic knowledge of the problem or a dedicated mathematical model of the production line setup. RL is model-free in the sense that the RL agent learns about its environment simply by interacting with it [22].

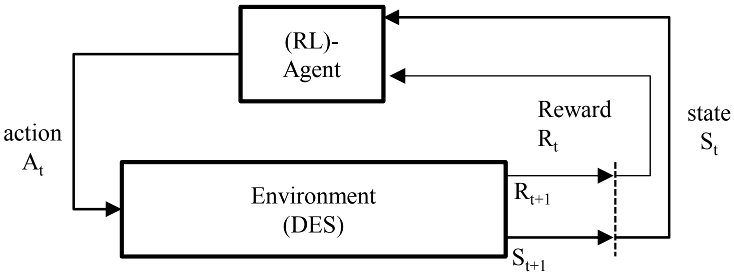


Figure 2: Markov Decision Process (MDP) as Default Framework for Reinforcement Learning (RL) [19]

RL can be understood as learning from a sequence of interactions between an agent and its environment, where the agent learns how to behave in order to achieve a goal [19]. The default formal framework to model RL problems is the Markov Decision Process (MDP), a sequential decision process modelled by a state space, an action space, and transition probabilities between states and rewards, see **Figure 2**. [18,22,19]

In an MDP, the agent acts based on observations of the states of the environment – in our case, these are the observations returned by the DES. The rewards received by the agent are the basis for evaluating these choices. The agent learns a policy resp. strategy, which may be understood as a function from state observations to actions. The agent's objective is to maximize the future cumulative discounted reward received over a sequence of actions [19]. This procedure is called *training*.

Especially the model-free character of reinforcement learning and the integration of simulations feedback data and thus the optimized parameter search motivate combining reinforcement learning with DES in this application. Previous work by the authors showed that this combination works in general and is very promising [13]. However, it is also stated that a promising practical application is unlikely without the explicit modelling of domain knowledge. Therefore, in the following paragraph an approach is presented, which describes the problem holistically based on a MDP and shows connecting points for the integration into a feasible improvement process for practice.

4. Methodology to improve production lines using reinforcement learning and simulation

To address the complex task of improving the economic performance of production lines, the task is formulated as a MDP for the approach presented here. This results in a problem structuring using a mathematical description, which remains comprehensible and application-oriented, since the problem is broken down into individual modules, which engineers can work with in a familiar manner based on their experience.

The basic idea of the approach is that an RL *agent* combines *actions* constrained by domain experts and sets explicit parameter values for them. For this purpose, the agent "plays" with the simulation model *(environment)* and learns from the observations of the returned *state* and the resulting monetary profit *(reward)* to choose reasonable improvement actions.

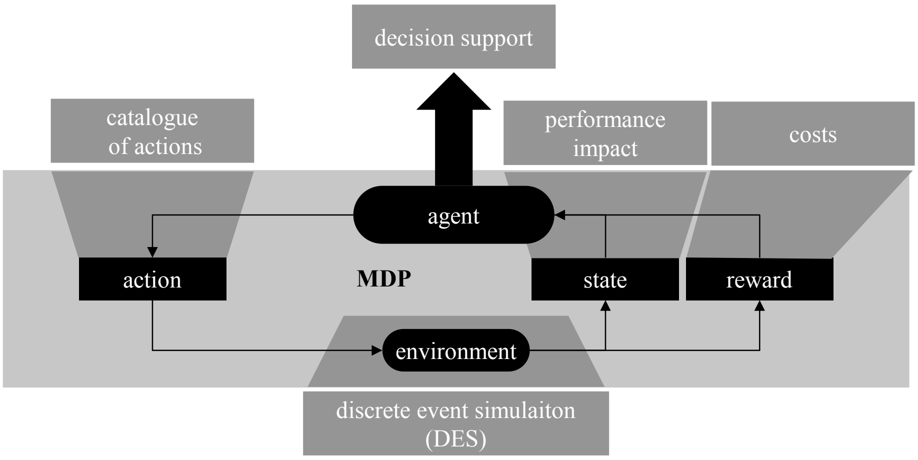


Figure 3: Approach interpreting the performance increase of production lines as MDP combining RL and DES

By evaluating these learnings of the agent (*policies)*, conclusions can be drawn about overall superior parameter combinations and their implementation order i.e. improvement trajectories to give decision support. **Figure 3** shows the interaction of these five modules and the translation into the corresponding formulation as a MDP. They are explained in more detail below.

Discrete event simulation (DES): Representation of the real production line taking into account the system complexity of production lines *(environment)*

The presented methodology requires a previously created and validated simulation model of a production line. A simulation model is necessary to represent the interdependence of the elements and thus the complexity of breakdowns in production lines. Methods for the data-based creation of such models can be found, for example, in. [34–36]. The presented approach uses standardized interfaces to communicate with established simulation software such as Siemens Plant Simulation [37].

Performance impact: Data based description of the influence of each element of a production line on the overall performance *(state)*

An explanatory model and method for measuring the influence of individual aggregates on the overall performance of a production line form the *performance impact* module. For this purpose, layout and aggregate information (main and auxiliary elements like conveyors) need to be linked with machine downtime and performance data. Layout information is collected through drawings, measurements and read-out data from machine controls and is already available in digital form in most companies of concern. Aggregate information is available in real resp. near time with modern production lines, since machine communication has become established through communication standards such as OMAC PackML [38]. An algorithm uses the combination of this data to allocate the performance losses to an individual aggregate taking into account the auxiliary and coupling elements. From breakdown of the losses per aggregate, the influence of each aggregate on the total performance of the production line can be determined. This influence on the performance of the production line from each individual aggregate makes it possible to prioritize based on the potential performance increase per aggregate, taking into account the performance losses in terms of OEE at constant machine speed. In terms of a MDP, this module of the method represents the observation of the current state and forms the state space. By integrating domain knowledge into the description of this observation, it is intended that the agent can identify potential process improvements more quickly and that learning time is reduced.

Catalogue of actions: Description of generic measures of performance increase for individual line elements *(action)*

In this module a catalogue of generic improvement measures as an explanatory model serves as a basis for later concretized improvement measures. For each of these measures, a description and abstraction in form of changes in simulation parameters and a cost function is created. This takes into account both investment and operating costs. The consideration of costs already at the stage of the creation of generic measures enables the later evaluation of the economic aspects of the improvement measures. On the basis of this catalogue, technically sensible and possible adjustment ranges for each parameter will be determined company and application specific to ensure practical applicability in individual settings. The description of this potential solution space forms the basis for the definition of the action space in terms of the MDP. Constraining the action space with expert and domain knowledge eliminates nonsensical parameter configurations, thus reducing the solution space and simplifying the agent's learning of effective strategies to improve the production line.

Costs: Evaluation of the concrete improvement measures in terms of economic benefit *(reward)*

This module evaluates the improvement measures chosen by the agent according to expected costs and potential performance improvement. For this, an equation is developed mapping the expected increase in performance and the expected costs of an improvement measure. The potential increase in performance is the result of the simulation. The costs are the result of the catalogue of actions. The development of this equation is based on traditional investment theory see e.g. [39,40]. Since different measures with different cost functions can change the same parameters of the production line, a heuristic is necessary to decide for the most economic measure for the parameter setting range of concern.

Decision support: Forming strategies for the economic performance increase of production lines in the form of improvement trajectories *(agent)*

The last module generates concrete action trajectories for improving production lines. These incremental steps can be interpreted by engineers and are practical and application-oriented, since adjustments to production lines in practice must also be made successively. For this purpose, the MDP descripted above is solved, which means an RL-agent learns to improve the system in terms of economic performance increase by maximizing the cumulative *reward*. Over time, the agent therefore identifies superior parameter combinations in the sense of sequencing individual measures at specific machines to trajectories. This is achieved by recording all parameter configurations tried throughout the training. The probability that a trajectory is superior is therefore higher for combinations of parameter configurations executed later, as the agent improves its strategies over time. Subsequently, these trajectories can be sorted according to the highest achieved reward and thus the most economical combinations are found. **Figure 4** shows such trajectories. The diameter of the circles in this figure represents the cumulative profit of the action. Thus, it can be seen that different trajectories can cause similar profits and that not every single change in the production system has to generate positive profit in isolation.

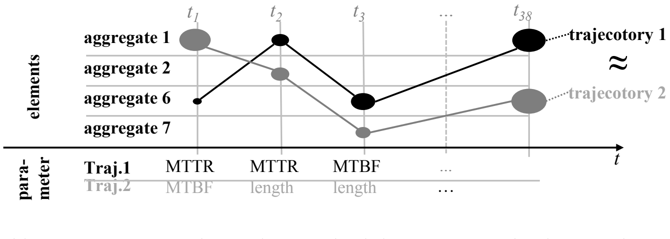


Figure 4: Visualisation of improvement trajectories as decision support for increasing performance of production lines

These combinations yield the basis for recommendations in terms of improvement trajectories. For these recommendations, a representation is developed which, in combination with a multi-criteria evaluation, provides the decision support for the practical modification of the production line.

5. Conclusion and further research

In this paper, a methodology for increasing the performance of production lines economically by identifiying alternative improvement trajectories using RL has been presented. The basic functionality has been proven by [13] and validated on an FMCG line. This paper embeds the problem solving method presented by [13] into a higher-level methodology for practical application.

Discussions of this approach with industrial enterprises continue to reveal a desire for a fixed budget for an optimization or improvement trajectory, which can be given to the RL-agent as additional constraint. In the detailed design it becomes apparent that the definition of the action space is critical for success and that the selection of variables by experts requires more precise support, since many engineers are not used to dealing with simulation-relevant parameters. A combination of the validation in [13] and this extended methodology is outstanding and is planned together with a comparison of different available algorithms as in [9].

Acknowledgement

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Bartkowiak, T., Ciszak, O., Jablonski, P., Myszkowski, A., Wisniewski, M., 2018. A Simulative Study Approach for Improving the Efficiency of Production Process of Floorboard Middle Layer, in: Advances in Manufacturing, Cham. 2018. Springer International Publishing, Cham, pp. 13–22.

[2] Bech, S., Brunoe, T.D., Nielsen, K., Andersen, A.-L., 2019. Product and Process Variety Management: Case study in the Food Industry. Procedia CIRP 81, 1065–1070.

[3] Nakajima, S., 1988. Introduction to TPM: Total productive maintenance. Productivity Press, Cambridge, Mass., 129 pp.

[4] Butollo, F., 2020. Digitalization and the geographies of production: Towards reshoring or global fragmentation? Competition & Change, 102452942091816.

[5] Andersson, C., Bellgran, M., 2015. On the complexity of using performance measures: Enhancing sustained production improvement capability by combining OEE and productivity. Journal of Manufacturing Systems 35, 144–154.

[6] Corrales, Lisbeth Del Carmen Ng, Lambán, M.P., Hernandez Korner, M.E., Royo, J., 2020. Overall Equipment Effectiveness: Systematic Literature Review and Overview of Different Approaches. Applied Sciences 10 (18), 6469.

[7] Koren, Y., Hu, S.J., Weber, T.W., 1998. Impact of Manufacturing System Configuration on Performance. CIRP Annals 47 (1), 369–372.

[8] Xi, S., Chen, Q., MacGregor Smith, J., Mao, N., Yu, A., Zhang, H., 2020. A new method for solving buffer allocation problem in large unbalanced production lines. International Journal of Production Research 58 (22), 6846–6867.

[9] Yegul, M.F., Erenay, F.S., Striepe, S., Yavuz, M., 2017. Improving configuration of complex production lines via simulation-based optimization. Computers & Industrial Engineering 109, 295–312.

[10] Rabe, M., Deininger, M., Juan, A.A., 2020. Speeding up computational times in simheuristics combining genetic algorithms with discrete-Event simulation. Simulation Modelling Practice and Theory 103, 102089.

[11] Karlsson, I., 2018. An interactive decision support system using simulation-based optimization and knowledge extraction: Dissertation Series. Doctoral thesis, monograph, Skövde, 80 pp.

[12] Trigueiro de Sousa Junior, W., Barra Montevechi, J.A., Carvalho Miranda, R. de, Teberga Campos, A., 2019. Discrete simulation-based optimization methods for industrial engineering problems: A systematic literature review. Computers & Industrial Engineering 128, 526–540.

[13] Schuh, G., Gützlaff, A., Schmidhuber, M., Maetschke, J., Barkhausen, M., Sivanesan, N., 2021. Identification of Superior Improvement Trajectories for Production Lines via Simulation-Based Optimization with Reinforcement Learning, in: Dolgui, A., Bernard, A., Lemoine, D., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems, vol. 634. Springer International Publishing, Cham, pp. 405–413.

[14] Wu, K., Zheng, M., Shen, Y., 2020. A generalization of the Theory of Constraints: Choosing the optimal improvement option with consideration of variability and costs. IISE Transactions 52 (3), 276–287.

[15] Ylipää, T., Skoogh, A., Bokrantz, J., Gopalakrishnan, M., 2017. Identification of maintenance improvement potential using OEE assessment. Int J Productivity & Perf Mgmt 66 (1), 126–143.

[16] Godinho Filho, M., Utiyama, M.H.R., 2015. Comparing different strategies for the allocation of improvement programmes in a flow shop environment. Int J Adv Manuf Technol 77 (5-8), 1365–1385.

[17] Burggräf, P., Wagner, J., Koke, B., Bamberg, M., 2020. Performance assessment methodology for AI-supported decision-making in production management. Procedia CIRP 93, 891–896.

[18] Gosavi, A. Solving Markov Decision Processes via Simulation: Handbook of simulation optimization, in: , vol. 216.

[19] Sutton, R.S., Barto, A., 2018. Reinforcement learning: An introduction, Second edition ed. The MIT Press, Cambridge, MA, London, 526 pp.

[20] Borrego, M., Foster, M.J., Froyd, J.E., 2014. Systematic Literature Reviews in Engineering Education and Other Developing Interdisciplinary Fields. J. Eng. Educ. 103 (1), 45–76.

[21] Bergeron, D., Jamali, M.A., Yamamoto, H., 2010. Modelling and analysis of manufacturing systems: a review of existing models. IJPD 10 (1/2/3), 46.

[22] Gosavi, A., 2015. Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning. Springer, New York 55. doi:10.1007/978-1-4899-7491-4, 508 pp.

[23] Liu, Y., Li, J., 2010. Split and merge production systems: performance analysis and structural properties. IIE Transactions 42 (6), 422–434.

[24] Nourelfath, M., Nahas, N., Ait-Kadi, D., 2005. Optimal design of series production lines with unreliable machines and finite buffers. J of Qual in Maintenance Eng 11 (2), 121–138.

[25] Tempelmeier, H., 2003. Practical considerations in the optimization of flow production systems. International Journal of Production Research 41 (1), 149–170.

[26] Spinellis, D.D., Papadopoulos, C.T., 2000. A simulated annealing approach for buffer allocation in reliable production lines. Annals of Operations Research 93 (1/4), 373–384.

[27] Jasiulewicz-Kaczmarek, M., Bartkowiak, T., 2016. Improving the performance of a filling line based on simulation. IOP Conf. Ser.: Mater. Sci. Eng. 145, 42024.

[28] Oljira, D.G., Abeya, T.G., Ofgera, G., Gopal, M., 2020. Manufacturing System Modeling and Performance Analysis of Mineral Water Production Line using ARENA Simulation. IJEAT 9 (5), 312–317.

[29] Umoren, I.U., Osueke, G.O., Okafor, B.E., 2021. Efficiency Analysis Associated with Production Line in Champion Breweries Plc. JMEA 11 (3).

[30] Jia, Z., Zhang, L., 2019. Serial production lines with geometric machines and finite production runs: performance analysis and system-theoretic properties. International Journal of Production Research 57 (8), 2247–2262.

[31] Yan, F.-Y., Wang, J.-Q., Li, Y., Cui, P.-H., 2021. An Improved Aggregation Method for Performance Analysis of Bernoulli Serial Production Lines. IEEE Trans. Automat. Sci. Eng. 18 (1), 114–121.

[32] Juan, A.A., Faulin, J., Grasman, S.E., Rabe, M., Figueira, G., 2015. A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. Operations Research Perspectives 2, 62–72.

[33] Hubscher-Younger, T., Mosterman, P.J., DeLand, S., Orqueda, O., Eastman, D., 2012. Integrating discrete-event and time-based models with optimization for resource allocation, in: 2012 Winter Simulation Conference. 2012 Winter Simulation Conference - (WSC 2012), Berlin, Germany. 12/9/2012 - 12/12/2012. IEEE, [Place of publication not identified], pp. 1–15.

[34] Gutenschwager, K., Rabe, M., Spieckermann, S., Wenzel, S., 2017. Simulation in Produktion und Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 290 pp.

[35] Rabe, M., Spieckermann, S., Wenzel, S., 2008. Verifikation und Validierung für die Simulation in Produktion und Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 242 pp.

[36] Vernickel, K., Brunner, L., Hoellthaler, G., Sansivieri, G., Härdtlein, C., Trauner, L., Bank, L., Fischer, J., Berg, J., 2020. Machine-Learning-Based Approach for Parameterizing Material Flow Simulation Models. Procedia CIRP 93, 407–412.

[37] Siemens. Use plant simulation and throughput optimization to improve manufacturing performance. https://www.plm.automation.siemens.com/global/de/products/manufacturing-planning/plant-simulation-throughput-optimization.html. Accessed 1 February 2022.

[38] OMAC, 2022. What is PackML? https://www.omac.org/packml. Accessed 1 February 2022.

[39] Hering, E. (Ed.), 2013. Taschenbuch für Wirtschaftsingenieure, 3., aktualisierte Auflage ed. Hanser Verlag, München, 634 pp.

[40] Römisch, P., Weiß, M., 2014. Projektierungspraxis Verarbeitungsanlagen. Springer Fachmedien Wiesbaden, Wiesbaden, 433 pp.

Biography

Günther Schuh (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University in Germany, is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT in Aachen and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen. Prof. Schuh is a member of several supervisory boards and boards of directors.

Andreas Gützlaff (*1989) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University in Germany. He is Chief Engineer and Head of Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Judith Fulterer (*1994) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University in Germany. She is a Research Assistant at Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University and Group Leader of the Production Logistics group in the Production Management department.

Jan Maetschke (*1994) studied Mechanical Engineering specializing in Production Engineering at RWTH Aachen University in Germany. He is a Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University and Group Leader of the Production Logistics group in the Production Management department.

3rd Conference on Production Systems and Logistics

A Bibliometric Analysis of Collaborative Supply Chain Risk Management in Crisis Situations

Ayman Nagi[1], Wolfgang Kersten[1]
[1]*Institute of Business Logistics and General Management, Hamburg University of Technology, Hamburg, Germany*

Abstract

Crises including the COVID-19 pandemic have caused disruptive changes to many industries and supply chains around the world. Their severe impacts on business and the economy provide an opportunity to increase preparedness and reveal the importance of implementing a collaborative supply chain risk management process. This paper uses a bibliometric analysis based on a co-citation analysis to reveal the research areas and gaps concerning collaborative supply chain risk management with a focus on crisis situations. Using a structured approach based on Soni and Kodali [1] and Gmür [2], 269 papers were extracted from the database Web of Science (WOS) using a specific search string. Data filtering and preparation using title, abstract, and full paper screening, as well as the number of cited-in references, led to a final sum of 50 papers. These papers were prepared for the co-citation analysis based on a co-citation matrix that served as an input for the Organizational Risk Analyzer (ORA) software. The cluster analysis was carried out in the ORA software with a threshold of 0.01, and based on that, five clusters were extracted from the network. Extracted main research areas include collaboration approaches and criteria as well as decision-making approaches and lessons learned from COVID-19. Research gaps and suggested future research areas are presented based on the clusters analysis.

Keywords

Bibliometric analysis; Crisis; Corona; COVID-19; Co-citation analysis; Supply Chain; Risk Management

1. Introduction

Due to the close interconnectedness of companies, interruptions and disruptions in the supply chain not only affect the acutely affected organization but can also result in financial losses and reputational damage for other organizations in the value network (see [3,4]). In addition, globalization increases supply chain complexity and makes supply and demand more volatile and difficult to forecast. The strong focus on efficiency in the context of supply chain management, which goes hand in hand with the reduction of buffer stocks in line with the lean philosophy, is increasingly making supply chains more vulnerable. Due to the high level of uncertainty and associated risks in global supply chains, it is of paramount importance for companies to understand the range of potential risks and their interconnectivity to establish appropriate risk mitigation strategies accordingly. These strategies should be accompanied by strong collaboration with supply chain partners to proactively manage different risk sources

Traditional supply chain risk management techniques rely on individual companies that define and implement mitigation measures for identified risks and their spillover effects. Therefore, collaborative

DOI: https://doi.org/10.15488/12181

approaches provide an opportunity to increase the effectiveness of the supply chain risk management process by focusing on interfirm relationship arrangements [5].

Based on the current COVID-19 pandemic, a large body of literature focuses on publishing empirical and theoretical studies for topics related, for example, to supply chain resilience and crisis management. According to the knowledge of the authors, no previous studies focused on conducting a bibliometric analysis for collaborative supply chain risk management with a focus on crisis situations. For this reason, this paper aims at analysing the body of literature in this regard based on a co-citation analysis to examine the research areas and gaps. Building on this, suggested future research areas are presented to tackle existing and potential crisis situations. The paper proceeds in Section 2 by providing a brief theoretical background concerning supply chain risk management, collaborative supply chain risk management, and bibliometric analysis. Afterwards, the methodology of the bibliometric analysis is elucidated in Section 3. Section 4 presents the results of the cluster analysis as well as the research areas, gaps, and suggested future research recommendations. Finally, Section 5 presents the conclusion and an outlook for further research.

2. Background

2.1 Supply Chain Risk Management

For a company to be optimally prepared against the risks that may arise and to minimize possible damage, a risk management system should be put in place. Risk management also exists in the supply chain; however, it differs from classic risk management. Particularly within value chains that operate globally and dynamically, comprehensive risk management is of crucial relevance [6]. Supply chain risk management (SCRM) is a developing research area, stemming from the growing recognition of the value of supply chain risk by practitioners and researchers [7]. In our understanding, supply chain risk management (SCRM) is to be understood as "[...] *a building block within supply chain management that encompasses all strategies and measures, all knowledge, all institutions, all processes, and all technologies that are suitable at the technical, personnel, and organisational levels for reducing risk within the supply chain.*" [8].

Proper assessment and planning using strategies, methods and tools for SCRM can minimize the impact of consequences that result from supply chain risks [9]. This necessitates a structured risk management process. The required steps comprise the identification, analysis, evaluation, and treatment of risks. Lastly, monitoring of risk management activities should be performed. These five steps of the risk management process should be carried out on a regular basis to meet the dynamic business environment [10].

Supply chain risk management is often linked with supply chain resilience. Supply chain resilience aims to maintain a certain desired performance in spite of disruptions. [11]. It is defined as "*the firm's capability to withstand, adapt, and recover from disruptions to meet customer demand, ensure target performance, and maintain operations in vulnerable environments*" [12]. Supply chain resilience implies not only the ability of a system to "bounce back" after a disrupting event but also the ability to adapt and transform [13].

2.2 Collaborative Supply Chain Risk Management

While coordination and collaboration are included in SCRM definitions, traditional SCRM approaches are not particularly effective in fostering inter-firm arrangements to deal with risk spillovers both within firms and across supply chains [5]. Collaboration among supply chain partners is the key mechanism for a good resilience against damage in case of any crisis [14] and can be categorized based on micro-, macro-and meso-levels [15].

The micro-level describes the direct coordination among organisations about supply risk prevention and recovery. The macro-level comes into place when organizations collaborate with other institutions such as the government, whereas the meso-level occurs when several supply networks work together on short- to medium-term supply risks [15]. Prerequisites for good cross-organisational collaboration are trust between the different actors, full traceability of the supply chain, awareness, knowledge of SCRM and its processes, and sharing of knowledge and information [14,16].

In seeking logistics solutions, risk managers embrace an attitude of exchange and collaboration with partners regarding aspects related to risk mitigation and sharing [17]. Risk information sharing, supplier trust, and shared SCRM understanding can influence the effectiveness of collaboration among supply chain partners [18]. Companies reject isolated practices and individualistic or opportunistic behaviours such as transferring and managing risk in isolation [17].

3. Methodology

Bibliometric analysis is a comparatively novel approach towards making sense of available metadata from a vast number of sources derived from e.g., scientific databases or search engines. Its purpose is to uncover emerging trends, collaboration patterns or explore the intellectual structure of a specific domain [19]. The bibliometric analysis differs from a systematic literature review (SLR) in the fact, that an SLR tend to rely on qualitative techniques and is far better suited for confined research areas and that a bibliometric analysis solely relies on quantitative analysis that reduces researcher and author bias [19]. Bibliometric analysis can handle a large number of literature sources compared to SLRs, which typically contain a smaller number of papers for review. To investigate data, bibliometric methods such as citation analysis, co-citation analysis and bibliographic coupling are normally used [20].

The co-citation analysis provides an effective methodology to analyse the relationship among core aspects of a specific scientific domain [21]. When two documents are cited together in one or more published articles, they are considered to be co-cited [22]. Co-citation count determines the proximity of content between two published articles [2]. This allows the extraction of clusters that correspond to research areas within a specific scientific domain. The overall research design of the performed analysis is adapted and further developed from [1] incorporating the bibliometric analysis based on [2] as illustrated in Figure 1.

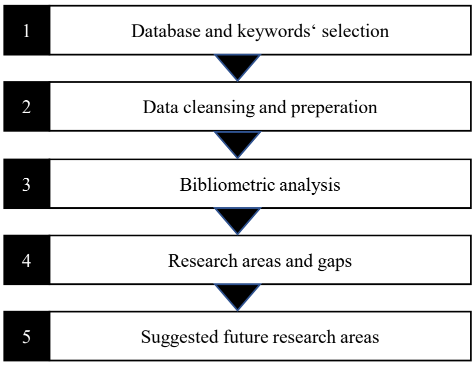


Figure 1: Approach of the research paper (own illustration based on [23])

First, Web of Science (WOS) was selected as a scientific database that includes required citation information. A specific set of keywords was selected as a search string to extract relevant articles from the database. This search string was entered in WOS using the logical operators “AND” and “OR” as follows:

- ("Cooperat*" OR "Co-operat*" OR "Collaborat*") AND "Risk manag*" AND "Crisis" (All Fields) **OR**
- ("Cooperat*" OR "Co-operat*" OR "Collaborat*") AND "Risk manag*" AND ("Corona" OR "Covid") (All Fields) **OR**
- ("Cooperat*" OR "Co-operat*" OR "Collaborat*") AND "supply chain" AND "Risk" AND ("Corona" OR "Covid") (All Fields)

Second, the data cleaning and preparation step mirrors the filtering procedure required to obtain a set of articles relevant to the area under study. Duplicates and articles with missing information were removed from the dataset, and the title and abstract of each article were then carefully read and reviewed to remove irrelevant articles. Afterwards, a full-text screening was conducted to extract the final set of articles for the descriptive and co-citation analysis.

Third, descriptive and co-citation analysis were conducted on the final set of articles, and the resulting findings were further examined in the fourth step to identify the main research areas and gaps in the current literature with regards to collaborative SCRM in crisis situations. Lastly, suggested future research areas based on the research gaps are proposed in the fifth step.

The adopted document co-citation approach in this study reflects the content proximity within a research discipline by analysing co-citations among selected peer-reviewed documents. Within the co-citation analysis, a specific method for calculating the co-citation frequencies is required as an input for the cluster analysis [2]. In this research, a Visual Basic for Applications (VBA) Macro code is developed in Microsoft (MS) Excel to calculate the co-citation frequencies. The cluster analysis is conducted using the Organisational Risk Analyser (ORA) software, a dynamic meta-network analysis and assessment tool developed by CASOS at Carnegie Mellon University in Pittsburgh, USA.

4. Results and Discussion

4.1 Descriptive analysis

The authors carefully selected 55 relevant articles from the 269 articles that emerged from the database queries based on the data cleaning and preparation phase. This phase comprises articles with missing information, title and abstract screening, as well as full-text screening. In total, 55 papers are considered to be relevant (see Figure 2).

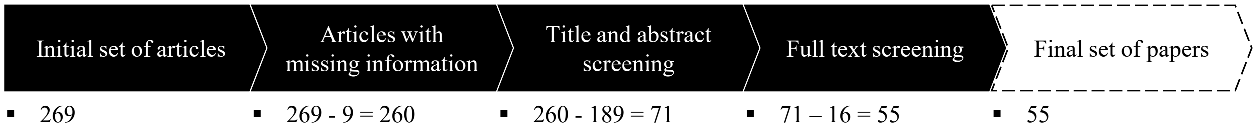


Figure 2: Data cleansing and preparation

In order to characterize this final dataset of relevant articles, descriptive figures are used. The descriptive analysis comprises the chronological development of the articles as well as the 10 most cited first authors.

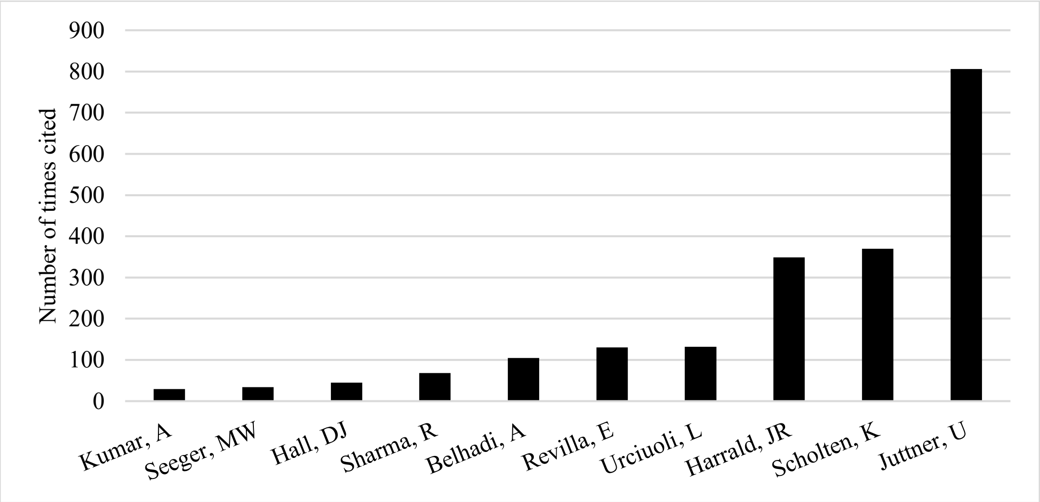


Figure 3: Most cited first authors

The most cited first author (Juttner, U), based on the extracted papers from WOS, focuses on supply chain resilience in the global financial crisis. Similarly, the author (Scholten, K) examined supply chain resilience and developed an integrated supply chain resilience framework. The third most cited first author (Harrald, JR) presented critical success factors in his paper to prepare and respond to extreme events. The other authors focus on different studies related for instance to risk mitigation strategies, resilience approaches, inter-organisational collaboration, as well as learnings from COVID-19.

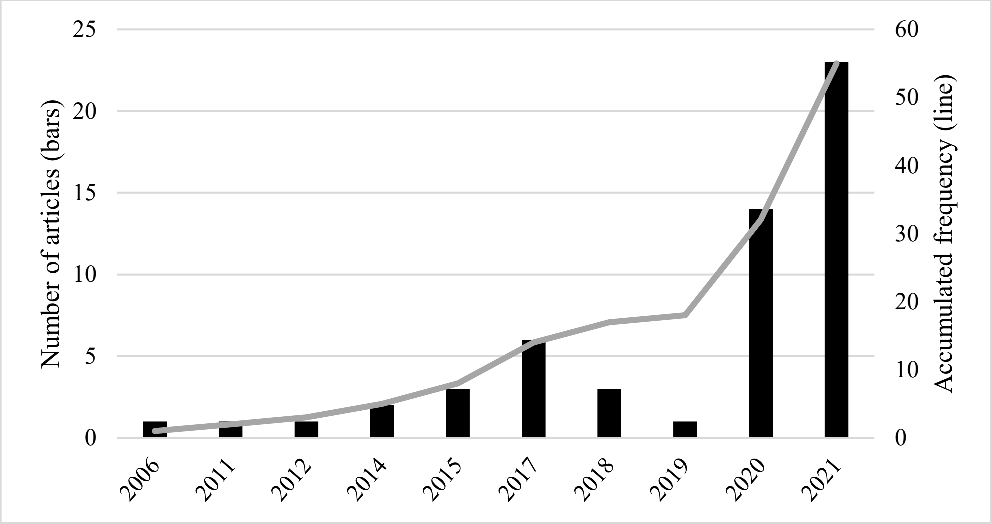


Figure 4: Number of papers per year

The chronological trend of the extracted papers is presented in Figure 4. An increase in the number of papers is observed in the years 2020 and 2021 due to the COVID-19 pandemic. Before that, the number of papers published for crises noticed a slight increase until 2017. There is a gap between 2006 and 2011 with no papers published that deal with collaboration in crisis situations. The cluster analysis of the final dataset is presented in the next subsection.

4.2 Cluster analysis

The results of the co-citation analysis are presented in this subsection based on the previously described steps (see Figure 5). All cited-in references for each article were stored separately in MS Excel worksheets that correspond to each article. Five papers from the dataset were not cited by any authors and therefore were excluded from the analysis. The cited-in references were downloaded using the library of Google Scholar and saved as CSV files for each article. In total, 2,455 cited-in references are distributed among the 50 articles.

For the calculation of the co-citation frequencies, a 50 x 50 raw co-citation matrix was programmed using VBA Macro in MS Excel. The co-citation matrix represents an integral input to the ORA software. Using a developed Macro, the raw co-citation matrix was generated by comparing the list of cited references for each article in each worksheet. By looping through each article, the Macro enters the frequency of co-citation in the appropriate field in the co-citation matrix.

The CoCit score was selected as the primary approach for creating the co-citation network and clusters. According to [24], the CoCit minimizes the relation of citation between the two co-citation partners. The approach adopts a value between 0 and 1 and associates the sum of co-citation counts with the mean and minimum values of the two individual citations.

The analysis was done in the ORA software with a threshold value of 0.01. This threshold value was adjusted manually until a clear pattern was detected. An additional revision of the articles' abstracts and introductions was conducted to extract the clusters from the network. Of the 29 articles in the final data set filtered using the threshold value, 20 are clustered references in the co-citation network.

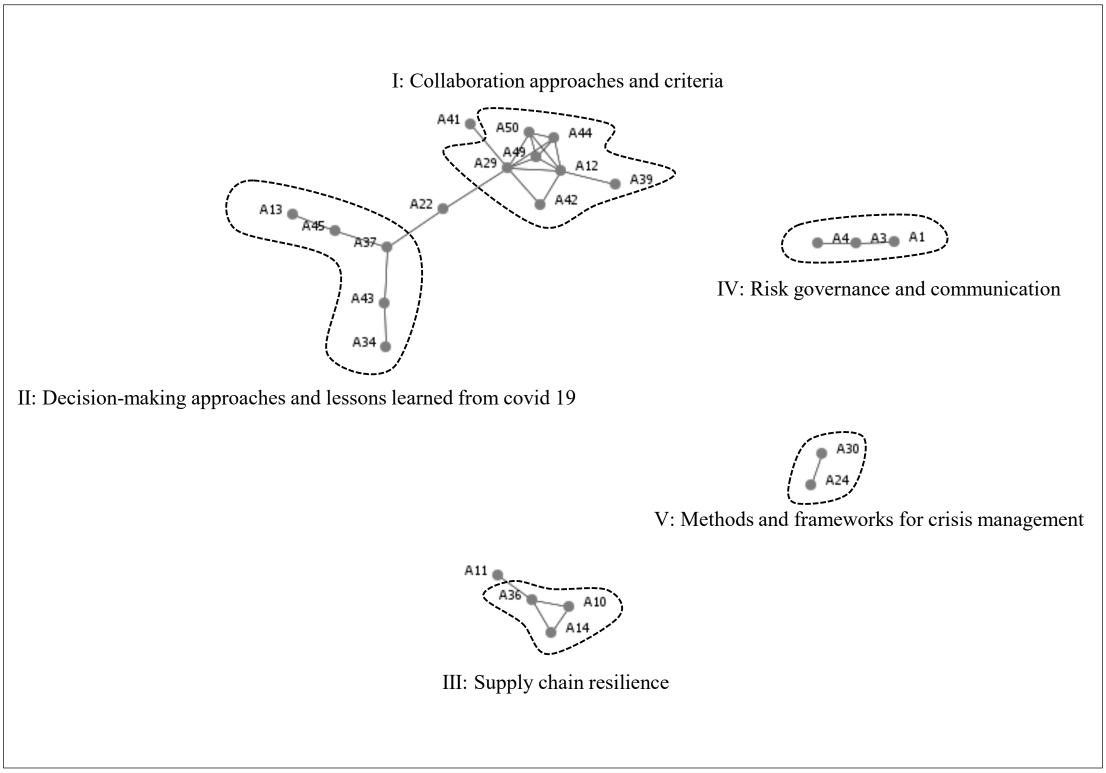


Figure 5: Clusters of the co-citation analysis

Cluster (I): Collaboration approaches and criteria (7 papers)

This cluster deals with collaborative approaches and criteria in different industries with a focus on the COVID-19 pandemic. Reference A42:[15] developed a typology of resiliency strategies concerning different collaboration types within and between supply networks. Reference A50:[25] analyse and discuss based on a literature review and case study research the relationships between fast-fashion retail chains and their suppliers' customers. The authors provided best business practices concerning cooperation with suppliers. Reference A12:[26] presents in their paper a collaborative approach for maintaining optimal inventory and mitigating stockout risks during a pandemic in healthcare supply chains based on a systematic literature review. Reference A39:[27] examine in their paper the role of relationship management between hotel chains and their key Tourism Supply Chain (TSC) agents to mitigate economic disruptions of epidemic outbreaks. Reference A49:[28] identified in their paper a total of 46 cross-sector collaboration activities based on the disaster management phases and resilience criteria that cover robustness, visibility, velocity and flexibility. Reference A29:[14] utilises seven semi-structured interviews with supply chain actors in the healthcare personal protective equipment supply chains as well as document analysis to analyse supply chain resilience during the COVID-19 pandemic response. The authors concluded that collaboration is considered a key to resilience. Reference A44:[29] found out in their paper that Collaboration Efficiency is the main criterion for accelerating the performance of Retail Supply Chains (RSCs) in a dynamic social environment. They concluded that RSCs require full integration and collaboration to mitigate the risks during and post-pandemic.

Cluster (II): Decision-making approaches and lessons learned from COVID-19 (5 papers)

This cluster deals with decision-making approaches for risk management and lessons learned from COVID-19. Reference A43:[30] present in their paper risk mitigation strategies for perishable food supply chains based on the fuzzy-best worst methodology (F-BWM). Reference A45:[31] focuses on the development of a framework to utilise lean, agile, and leagile strategies in the supply chains. The authors analysed as well the impact of these strategies on crisis using the example of COVID-19. Reference A34:[32] examined agricultural supply chains risk caused by disruptions and identified strategies for decision-makers such as supply chain collaboration and shared responsibility. Reference A13:[16] stress the importance of involving communities in decisions during and after a crisis event occurs. The authors propose that risk managers may benefit from incorporating collaborative planning principles in their approaches, especially at the prevention stage. Using seven companies from different industries, supply chain positions, and countries, reference A37:[33] examine how insights from theories of the total cost of ownership, supplier segmentation, and supply chain change management can be applied to efforts to manage COVID-19 risks and disruptions in the supply chain.

Cluster (III): Supply Chain Resilience (3 papers)

This cluster comprises three papers that deal with supply chain resilience. Reference A10:[34] conceptualizes supply chain resilience and investigate its related concepts of SCRM and supply chain vulnerability. The authors of A36:[35] develop an integrated supply chain resilience framework utilising a qualitative case of a collaborative agency. Finally, reference A14:[36] analyses how energy supply chains function to increase resilience in the face of exogenous security threats and what support mechanisms the European Union should subsequently introduce or improve.

Cluster (IV): Risk governance and communication (3 papers)

This cluster deals with studies related to risk governance and communication with stakeholders. The authors of A4:[37] focus on defining the term supply chain governance and developed an associated conceptual framework that reflects different types of supply chains and actors. Reference A1:[38] deals with risk information sharing and investigates communication challenges linked to risk and vulnerability assessment. Similarly, reference A3:[39] focuses on communicating risk in disaster management systems, and based on two experiments, the authors reached a conclusion that the presence of risk information greatly influences the ability of stakeholders to carry out well-informed decisions.

Cluster (V): Methods and frameworks for crisis management (2 papers)

This small cluster consists of two papers that present methods and frameworks for crisis management. Reference A24:[40] developed a method that integrates Business Impact Assessment (BIA) and Risk and Vulnerability Assessment (RVA) for the public crisis management sector. Reference A30:[41] develops a multilevel framework to enhance organisational resilience for responding to crises. The authors argue that crisis management and organisational resilience are shaped mutually across different levels, from environmental, organisational, to individual.

4.3 Main Research areas and gaps

Three main research areas were extracted based on the frequency of articles in each research area from the co-citation network. The threshold to detect a main research area is set to be three articles. First, collaborative approaches and criteria for different industries are thoroughly analysed by different authors as can be observed in the first cluster. For instance, papers related to this area investigate collaborative approaches for inventory optimization and criteria such as collaboration efficiency. Second, there is a focus on resilience as well as supply chain resilience approaches and frameworks that were developed for crisis management which can be noticed particularly in the third and fifth clusters. Papers from these clusters utilized conceptual analysis, qualitative case study as well as empirical studies. Third, decision-making approaches and lessons learned from the COVID-19 pandemic (see cluster II) is a main research area that tackles supply chain risks caused by disruptions. The incorporation of communities and collaborative planning principles are examples of decision-making approaches from this cluster to manage supply chain risks caused by disruptions.

From the cluster analysis, only a few studies focused on the role of risk governance and risk communication in managing risks or crisis situations. Additionally, only a few papers developed methods and frameworks for crisis situations based on SCRM. There is a lack of papers that provides conceptual analysis and a roadmap for implementing collaborative SCRM with a focus on crisis situations. None of the papers as well considered the integration of business continuity management with collaborative SCRM. Based on the aforementioned research gaps, suggested future research areas are elaborated in the next subsection.

4.4 Suggested future research areas

There is a need to conduct further research concerning frameworks and models that can guide companies in understanding the requirements for implementing a collaborative SCRM process. Empirical studies based on interview and survey studies can examine the current status of collaborative SCRM and extract implementation aspects. There is also a need to analyse the impact of risk governance on collaborative SCRM. In this regard, case studies, as well as explorative approaches, are recommended to understand the current situation, challenges, and opportunities for collaborative risk management. Transdisciplinary studies integrating related research fields such as resilience and business continuity management are recommended to develop holistic frameworks and models that support collaboration aspects, especially in crisis situations. Studies that define maturity levels linked to Key Performance Indicators (KPI) for collaborative performance systems can help companies to understand and improve their current collaborative risk management level (see [5]). An operationalization process is required in advance to enable the proper assessment of collaboration in SCRM.

5. Conclusion and outlook

This paper utilised a bibliometric analysis based on a co-citation analysis to reveal the research areas and gaps concerning collaborative SCRM with a focus on crisis situations. Based on the analysis, three main research areas are extracted: (1) collaborative approaches and criteria for different industries such as healthcare and fashion (2) resilience and supply chain resilience approaches and frameworks for crisis management (3) decision-making approaches and lessons learned from the COVID-19 pandemic. Besides the research areas, the research gaps are extracted based on the cluster analysis. A gap was detected concerning methods and frameworks for crisis situations based on SCRM. Another deficiency is connected to studies that provide conceptual analysis and a roadmap for developing a collaborative SCRM with a focus on crisis situations. Lastly, a clear gap is noticed with regards to the integration of business continuity management with collaborative SCRM. Based on the research gaps, future research areas are suggested covering collaborative SCRM, business continuity management, resilience and risk governance encompassing theoretical, conceptual, and explorative approaches.

The co-citation analysis performed in this study has several limitations. First, the extracted papers were based on a specific search string that could have omitted other relevant papers. Second, several papers were not cited at all or only cited by a few authors since a large number of papers were published in the years 2020 and 2021. Third, the cluster analysis was based on the CoCit method for generating the co-citation network. Future studies should consider applying a Multi Vocal Literature Review (MLR) to systematically analyse both white and grey papers. The current research indicates a clear research gap concerning holistic frameworks and models for implementing collaborative SCRM. Therefore, it is recommended to develop theoretical and conceptual frameworks as well as models that present the building blocks and aspects for implementing a collaborative SCRM from theory and practice. These models and frameworks should investigate, for instance, the role of supply chain risk governance on collaborative SCRM as well as investigate how collaboration approaches for SCRM affect crisis management. Operationalization and quantification approaches that measure and assess the successful implementation of a collaborative SCRM and the intensity of collaboration should be examined in further research. The next step in our research is to develop a conceptual framework for collaborative supply chain risk management with a focus on crisis situations.

Acknowledgements

This research is supported by the Authority for Science, Research, Equality and Districts (BWFGB) based on the research project: “Entwicklung einer Online-Wissensplattform für unternehmensübergreifendes Supply Chain Risikomanagement in Krisensituationen” EN: “Development of an online knowledge platform for cross-company SCRM in crisis situations”.

Biography

Ayman Nagi is a research associate in the Institute of Business Logistics and General Management at the Hamburg University of Technology. His research is focusing on Risk Management in seaports and collaborative Supply Chain Risk Management. Previously, he has worked in the fields of Quality, Production and Risk Management with several international and national companies in Germany and Jordan. He received his M.Sc. in International Production Management from the Hamburg University of Technology (Germany) and his MBA in Technology Management from the Northern Institute of Technology Management (Germany) in 2017.

Wolfgang Kersten is a full professor and head of the Institute of Logistics and General Management at the Hamburg University of Technology (Germany). His research focuses on the Digital Transformation of Logistics as well as Complexity, Risk and Sustainability Management of value chains. He received his diploma in Industrial Engineering from the Technical University of Darmstadt (Germany) and his PhD from the University of Passau (Germany). After various management positions at Mercedes Benz AG, he was appointed to the Chair of Production Economics at the Hamburg University of Technology in 1998.

References

[1] Soni, G., Kodali, R., 2011. A critical analysis of supply chain management content in empirical research. Business Process Mgmt Journal 17 (2), 238–266.

[2] Gmür, M., 2003. Co-citation analysis and the search for invisible colleges: A methodological evaluation. Scientometrics 57 (1), 27–57.

[3] Norrman, A., Wieland, A., 2020. The development of supply chain risk management over time: revisiting Ericsson. IJPDLM 50 (6), 641–666.

[4] Paul, S.K., Asian, S., Goh, M., Torabi, S.A., 2019. Managing sudden transportation disruptions in supply chains under delivery delay and quantity loss. Ann Oper Res 273 (1-2), 783–814.

[5] Friday, D., Ryan, S., Sridharan, R., Collins, D., 2018. Collaborative risk management: a systematic literature review. IJPDLM 48 (3), 231–253.

[6] Kersten, W., Hohrath, P., Winter, M., 2008. Risikomanagement in Wertschöpfungsnetzwerken–Status quo und aktuelle Herausforderungen. Supply Chain Risk Management 7, 7–22.

[7] Sodhi, M.S., Son, B.-G., Tang, C.S., 2012. Researchers' perspectives on supply chain risk management. Production and operations management 21 (1), 1–13.

[8] Kersten, W., Held, T., Meyer, C., Hohrath, P., 2007. Komplexitäts-und Risikomanagement als Methodenbausteine des Supply Chain Managements, in: Wildemann, H. (Ed.), Management am Puls der Zeit-Strategien, Konzepte und Methoden. TCW Transfer-Centrum, München.

[9] Gurtu, A., Johny, J., 2021. Supply Chain Risk Management: Literature Review. Risks 9 (1), 16.

[10] Kersten, W., Schroeder, M., Nagi, A., in press. Digitalisation - A challenging enabler for supply chain risk management, in: Roth, S., Corsten, H. (Eds.), Handbuch Digitalisierung. Vahlen.

[11] Ivanov, D., Tsipoulanidis, A., Schönberger, J., 2021. Supply Chain Risk Management and Resilience, in: Ivanov, D., Tsipoulanidis, A., Schönberger, J. (Eds.), Global Supply Chain and Operations Management. Springer International Publishing, Cham, pp. 485–520.

[12] Hosseini, S., Ivanov, D., Dolgui, A., 2019. Review of quantitative methods for supply chain resilience analysis. Logistics and Transportation Review 125, 285–307.

[13] Wieland, A., Durach, C.F., 2021. Two perspectives on supply chain resilience. J Bus Logist 42 (3), 315–322.

[14] Scala, B., Lindsay, C.F., 2021. Supply chain resilience during pandemic disruption: evidence from healthcare. SCM 26 (6), 672–688.

[15] Azadegan, A., Dooley, K., 2021. A Typology of Supply Network Resilience Strategies: Complex Collaborations in a Complex World. J Supply Chain Manag 57 (1), 17–26.

[16] Shmueli, D.F., Ozawa, C.P., Kaufman, S., 2021. Collaborative planning principles for disaster preparedness. International Journal of Disaster Risk Reduction 52 (1), 1–8.

[17] Lavastre, O., Gunasekaran, A., Spalanzani, A., 2012. Supply chain risk management in French companies. Decision Support Systems 52 (4), 828–838.

[18] Gang Li, Huan Fan, Peter K.C. Lee, T.C.E. Cheng, 2015. Joint supply chain risk management: An agency and collaboration perspective. International Journal of Production Economics 164 (1), 83–94.

[19] Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., Lim, W.M., 2021. How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research 133 (1), 285–296.

[20] Wallin, J.A., 2005. Bibliometric methods: pitfalls and possibilities. Basic & clinical pharmacology & toxicology 97 (5), 261–275.

[21] Ferreira, J.J.M., Fernandes, C.I., Ratten, V., 2016. A co-citation bibliometric analysis of strategic management research. Scientometrics 109 (1), 1–32.

[22] Smith, L.C., 1981. Citation analysis. Bibliometrics, Library Trends 30 (1), 83–106.

[23] Nagi, A., Indorf, M., Kersten, W., 2017. Bibliometric analysis of risk management in seaports, in: Hamburg International Conference of Logistics (HICL) 2017, 491–521.

[24] Gmür, M., 2003. Co-citation analysis and the search for invisible colleges: A methodological evaluation. Scientometrics 57 (1), 27–57.

[25] Dewalska-Opitek, A., Bilińska-Reformat, K., 2021. To What Extent Retail Chains' Relationships with Suppliers Make the Business Trustworthy—An Empirical Study on Fast Fashion in Pandemic Times. Journal of Risk and Financial Management 14 (4), 1–11.

[26] Friday, D., Savage, D.A., Melnyk, S.A., Harrison, N., Ryan, S., Wechtler, H., 2021. A collaborative approach to maintaining optimal inventory and mitigating stockout risks during a pandemic: capabilities for enabling health-care supply chain resilience. Journal of Humanitarian Logistics and Supply Chain Management 11 (2), 248–271.

[27] González-Torres, T., Rodríguéz-Sánchez, J.-L., Pelechano-Barahona, E., 2021. Managing relationships in the Tourism Supply Chain to overcome epidemic outbreaks: The case of COVID-19 and the hospitality industry in Spain. International Journal of Hospitality Management 92, 1–11.

[28] Medel, K., Kousar, R., Masood, T., 2020. A collaboration–resilience framework for disaster management supply networks: a case study of the Philippines. JHLSCM 10 (4), 509–553.

[29] Sharma, M., Luthra, S., Joshi, S., Kumar, A., 2021. Accelerating retail supply chain performance against pandemic disruption: adopting resilient strategies to mitigate the long-term effects. Journal of Enterprise Information Management 34 (6), 1844–1873.

[30] Kumar, A., Mangla, S.K., Kumar, P., Song, M., 2021. Mitigate risks in perishable food supply chains: Learning from COVID-19. Technological Forecasting and Social Change 166, 120643.

[31] Rashad, W., Nedelko, Z., 2020. Global Sourcing Strategies: A Framework for Lean, Agile, and Leagile. Sustainability 12 (17), 7199.

[32] Sharma, R., Shishodia, A., Kamble, S., Gunasekaran, A., Belhadi, A., 2020. Agriculture supply chain risks and COVID-19: mitigation strategies and implications for the practitioners. International Journal of Logistics Research and Applications, 1–27.

[33] van Hoek, R., 2020. Responding to COVID-19 Supply Chain Risks—Insights from Supply Chain Change Management, Total Cost of Ownership and Supplier Segmentation Theory. Logistics 4 (4), 23.

[34] Jüttner, U., Maklan, S., 2011. Supply chain resilience in the global financial crisis: an empirical study. Supply Chain Management: An International Journal 16 (4), 246–259.

[35] Scholten, K., Sharkey Scott, P., Fynes, B., 2014. Mitigation processes – antecedents for building supply chain resilience. SCM 19 (2), 211–228.

[36] Urciuoli, L., Mohanty, S., Hintsa, J., Gerine Boekesteijn, E., 2014. The resilience of energy supply chains: a multiple case study approach on oil and gas supply chains to Europe. Supply Chain Management: An International Journal 19 (1), 46–63.

[37] Ahlqvist, V., Norrman, A., Jahre, M., 2020. Supply Chain Risk Governance: Towards a Conceptual Multi-Level Framework. OSCM: An Int. Journal 13 (4), 382–395.

[38] Lin, L., Abrahamsson, M., 2015. Communicational challenges in disaster risk management: Risk information sharing and stakeholder collaboration through risk and vulnerability assessments in Sweden. Risk Manag 17 (3), 165–178.

[39] Lin, L., Rivera, C., Abrahamsson, M., Tehler, H., 2017. Communicating risk in disaster risk management systems – experimental evidence of the perceived usefulness of risk descriptions. Journal of Risk Research 20 (12), 1534–1553.

[40] Hassel, H., Cedergren, A., 2021. Integrating risk assessment and business impact assessment in the public crisis management sector. International Journal of Disaster Risk Reduction 56 (1), 1–14.

[41] Tasic, J., Amir, S., Tan, J., Khader, M., 2020. A multilevel framework to enhance organizational resilience. Journal of Risk Research 23 (6), 713–738.

3rd Conference on Production Systems and Logistics

Supporting The Transformation To Climate Neutral Production With Shop Floor Management

Sebastian Bardy [1], Stefan Seyfried[1], Joachim Metternich [1], Matthias Weigold[1]

[1]*Institute of Production Management, Technology and Machine Tools, Otto-Berndt-Straße 2, 64287 Darmstadt, Germany*

Abstract

The European Green Deal proposes the transformation to climate neutrality by 2050. Especially for manufacturing companies and their production sites, this transformation is a big challenge. Every aspect of the value stream needs to be re-evaluated and adjusted to reach the new target state of climate neutral production. In the last decades, many companies used lean management methods to improve production in the dimensions of time, quality, and cost. However, a growing number of studies show that lean methods can also be used to drive sustainability goals (with the target state being climate neutral production). This paper analyses the suitability of shop floor management, a popular lean method, in the context of climate neutral production. To this end, a literature research has been conducted to summarize the goals of shop floor management and the success factors for the transformation to climate neutral production. Then the results are contrasted and overlaps are analysed to identify possible shop floor management tools to accelerate the transformation to climate neutral production. Finally, the findings are briefly discussed and summarized in a matrix. The paper closes with specific recommendations for further research in this area.

Keywords

shop floor management; lean management; climate neutral production; climate neutrality; sustainable production

1. Motivation

With the Paris Agreement the United Nations agreed that global warming is to be limited to a maximum of 2 °C and that efforts are to be pursued to remain below 1.5 °C [1]. It was recently reaffirmed at COP26 in Glasgow [2]. This goal requires steps to be taken in all areas of society including industrial production. The manufacturing sector is challenged to integrate the goal of climate neutrality into operational company processes. In the last decades many companies used lean management methods to improve production in the dimensions of time, quality, and cost. An extension of the dimensions is therefore plausible and has already been investigated by several studies [3–5]. The difficulty is that elimination of waste (lean) and resource efficiency (climate neutrality) are dealt with in different committees or working groups [3]. Furthermore, approaches to Lean&Green to date are project-based and still focus on cost reduction as the main dimension [6]. For climate neutral production a much more fundamental and sustainable change in the companies must be achieved [7]. In their lead research study the "Deutsche Energie-Agentur" (dena) describes the change necessary in vision and strategic goals of companies to reach climate neutrality [8]. To operationalise strategic goals the widely accepted lean method of shop floor management (SFM) can be used. In 2018, a study from Germany shows that more than 80 percent of producing companies use SFM to lead workers on the shop floor and help them improve production processes [9]. Since climate neutrality is also a strategic

DOI: https://doi.org/10.15488/12164

improvement process the question is whether SFM can support the success factors of climate neutral production. However, there is currently a lack of a comprehensive overview of the success factors for climate neutrality in production. Based on that, a comparison can be made to develop recommendations in which areas SFM can promote climate-neutral production.

This paper is divided into three parts. First, the state of the art for both areas is analysed. The second part explores the success factors of climate neutral production through a systematic literature review. Then, matching areas are analysed and recommendations for further research are discussed.

2. State of the art

2.1 Climate neutral production

In recent years, climate targets of various countries and other institutions have been tightened. The goals are often defined by the target state of climate neutrality. The European Union, for example, has set a target in the European Green Deal 2019 to become the first climate neutral continent by reducing net greenhouse gas emissions to zero by 2050 [10] and by 55 percent by 2035 compared to emissions in 1990 [11].

The term climate neutrality must be distinguished from greenhouse gas neutrality and CO_2 neutrality. CO_2 neutrality refers to a state in which as much CO_2 is emitted as is absorbed by natural and technical sinks. Additionally, greenhouse gas neutrality includes other greenhouse gases. Finally, climate neutrality is the condition in which no further increase of global temperatures exists. Since there are many effects on climate change, some of which are not directly measurable, the term climate neutrality is often used synonymously with greenhouse gas neutrality [7,12]. Despite significant reductions in recent years, the industrial sector still accounts for about 20 percent of greenhouse gas emissions of the European Union [10]. Accordingly, there is still a need for action in the industrial sector to further reduce emissions and contribute to the goal of climate neutrality. In addition to the ecological impact the goal of climate neutrality is also gaining importance for companies due to other reasons: The climate targets are becoming an important criterion for investors to base their decisions on. Emission certificates must be purchased, whereby rising prices can be assumed. Rising energy prices also make activities to increase energy efficiency economically attractive. In addition, the demand for climate neutral products on the customer side is growing as well. [13]

A company's emissions are often broken down into three scopes according to the Greenhouse Gas Protocol [14]. Scope 1 includes direct emissions inside the company e.g., from the combustion of natural gas. Scope 2 covers emissions from energy purchases. Scope 3 includes further emissions, for example caused by upstream and downstream processes in the supply chain and the use of the product [15]. In all three scopes, there is a multitude of possible actions for companies to reduce emissions.

The goal of climate neutrality is often examined in cross-sectoral transformation studies for entire national economies [8,16,17]. In addition, there are studies that investigate transformation paths in specific sectors, such as the industrial sector, or specific industries [18]. Nevertheless, for a comparison of shop floor management with climate neutral production, an overview of success factors for climate neutrality is missing. There are guidelines for implementation, but these focus on partial aspects such as resource efficiency [19] or are limited to certain stakeholders, such as managers [8]. Therefore, an overview of the success factors must be created, on which further considerations will be built.

2.2 Shop floor management

SFM originated in lean manufacturing theories and was first introduced by Suzaki in 1993 with the book "The New Shop Floor Management" [20]. On the shop floor goals set by the management are translated into traceable key performance indicators (KPIs). SFM helps to identify deviations in the KPIs, analyse those in shop floor meetings and initiate a problem-solving process. Optimizations developed in the problem-solving process are stabilized and standardized in order to reach a continuous improvement of the production processes [20,21]. The Darmstadt SFM model developed by Hertle et. al. identifies five core elements of shop floor management: performance management, problem management, lean leadership, glass wall management and competence management. For each element several goals and methods are presented in the following paragraphs:

With **performance management** the work scheduling as well as the order scheduling in production is addressed [20–22]. In performance management, a basic distinction is made between two types of improvements: Reacting to problems and achieving more challenging goals [20–22]. A major lean principle for the reaction to problems is the *stabilization of processes* [20,21]. The initial standard is the baseline for further improvement and is controlled through KPIs. If there are regular deviations generated by unstable processes, these must first be addressed [23]. The *definition of goals* for the production processes by management is the prerequisite for the *definition of KPIs* to visualize the degree of target achievement. This goal oriented approach to SFM is linked with the lean method "Hoshin Kanri", where every KPI or goal is linked to the long-time vision of the company [24]. All entities in the company are encouraged to have the same structure of KPIs to strengthen transparency [25]. KPIs are challenged every day and optimized towards the set goals to *control and increase the production output* [21,23].

In **problem management** the deviations in KPIs are presented to the employees and managers where they must make the decision how to handle the deviation. This can be categorized into three different reactions: If the deviation is not impactful or even a false alarm it can be ignored. If the deviation has an impact on production performance immediate action to prevent further damage or solve minor issues are to be taken. Finally, if the deviation is a major problem a *systematic problem-solving process* (SPSP) is used to find the right countermeasure. [20,21,24]

Lean leadership is strongly linked with the "Genchi Genbutsu" method of lean management [26]. Managers do not spend enough time on the shopfloor and loose connection to the problems of their employees [23,27]. By meeting on the shop floor together with their employees (Go and See), problems can be analysed and addressed faster. To further strengthen this strategy the hierarchical depth of management is reduced and managers are encouraged to be active on the shop floor to see for themselves. [28,29]

The reduction of hierarchical depth also implies a reorganization of the work environment into smaller teams. Through the introduction of mini factories for *visual management*, the teams are responsible for their own area. Employees start to identify themselves with their work and communication with other mini factories is strengthened [29]. Such communication is only possible by the use of tools for transparency like shop floor boards or Andon that enable to analyse production efficiency even as an external visitor [30,31]. In addition, **glass wall management** includes the goal of *information dissemination*. In order to empower employees there is a need for systematic information dissemination from management to their employees [28]. The employee needs to know the goals/vision created by managers to understand their decisions [29].

In **competence management** three goals are identified. As described beforehand SFM involves employees in the continuous improvement process. This *transfer of responsibility to employees* can be seen as an empowerment to carry out adjustments/optimizations of processes independently/together with colleagues [28,32]. Shop floor meetings are held each day at a set time with a series of topics that are discussed in a predefined order. This *structured communication* ensures that all relevant topics are discussed in a set timeframe while still leaving room for open discussion [23]. Moreover, the transfer of responsibility and the structured communication are used to achieve *competence development for employees as well as management*. Competence development on all hierarchical levels is necessary to have constant continuous improvement for the production process [28]. To achieve competence development, employees as well as management are given specific roles such as shop floor operator, manager and team leader [23]. Leading personnel must be developed into methodical coaches willing to set an example for their employees[23,28]. Furthermore, targeted, work-integrated competence development is key for employees [28].

3. Systematic literature review to identify success factors for climate neutral production

To identify success factors for climate neutral production a systematic literature research was conducted on the database of ScienceDirect. Only literature in English was considered. The database of ScienceDirect was scanned with the following search string: (("climate neutral" OR "carbon neutral") OR ("climate neutrality" OR "carbon neutrality" OR "net-zero")) AND ("manufacturing" OR "industrial production" OR "industry"). The query was limited to title, abstract or author-specified keywords. The focus is on articles that show the implications of the goal of climate neutrality for industrial production, and less on the description of specific technologies. From the results of the literature research, 16 success factors for climate neutral production were derived (see Table 1). Contrasting the methods of SFM these were grouped according to technical, organisational and human success factors [33]. The success factors are chosen to ensure that they can be actively influenced by a manufacturing company. External influencing factors such as the development of the regulatory framework or the energy system outside the factory are just as relevant for achieving climate neutrality but are not within a company's direct sphere of influence. They are therefore neglected in this analysis. Requirements for the implementation of the success factors are included in the description.

Table 1: Success factors of climate neutral production

Technical success factor	**Description**
Increasing energy efficiency	*Increasing energy efficiency* is mentioned in a variety of studies across industry as a key success factor for the development towards climate neutrality [8,16,34–39]. The implementation of energy efficiency measures is also of importance because it supports the implementation of other success factors by reducing the overall energy consumption of the industry sector. Energy efficiency measures need to be targeted after identifying potentials for increasing energy efficiency. They can range from improvements in specific machines and processes to complex systemic measures [39]. Besides the ecological effect, energy efficiency measures can contribute to the profitability of a company by reducing energy costs [40].
Increasing energy flexibility	Energy flexibility describes the ability of a system to adapt to changes in the energy market, especially by shifting electrical loads over time[41,42]. Hence, *increasing energy flexibility* on the demand side creates the prerequisite for integrating an increased percentage of volatile renewable energy sources into the energy system. [8,16] Thus, energy flexibility is necessary not only for transforming factories but also the surrounding energy system towards climate neutrality. *Increasing energy flexibility* can be achieved for example by using energy storages or the planned adaption of production processes according to energy forecasts. [41]
Increasing material efficiency	*Increasing material efficiency* contributes to reducing material-related greenhouse gas emissions by efficient product design and reducing the creation of waste during production by ensuring high quality production processes. [8,16,37–40]. The increase of material efficiency as a success factor includes material substitution as well [39].
Reuse of waste heat	The *reuse of waste heat* can be seen as a sub-aspect of increasing energy efficiency. However, it is of such importance for climate neutrality to decarbonise the heat supply that it is listed as a separate success factor. [8,36,43]
Electrification of energy demands	*Electrification of energy demands* can help reduce direct emissions at Scope 1 by replacing fossil fuels [8,12,37–39,44–46]. Heat pumps are a well-known example of those power-to-heat technologies [38].
Using green fuels	*Using green fuels* covers synthetic fuels, hydrogen, and biomass. These fuels can help replace fossil fuels [8,12,16,36–39,43,44,46,47]. The emission reduction is based on the use of electricity to produce synthetic fuels or hydrogen, or the use of biological processes in the case of biomass. Their use is indicated when direct electrification is not an option [16,45].
Using renewable energies	*Using renewable energies* is an obvious success factor and a prerequisite for other success factors such as electrification or the use of green fuels. Companies can both switch their energy purchases to renewable energies and operate renewable energy plants themselves. [8,16,34,35,43,48]

Implementing elements of circular economy	*Implementing elements of circular economy* contributes to reducing the footprint of the used materials and the manufactured goods. [8,16,37–40,44]
Carbon capture and storage/usage (CCS/CCU)	*Carbon capture and storage (CCS)* processes may be used to avoid emissions, for example when the application of the technical success factors described above does not lead to achieving complete climate neutrality [37–39,48–51]. In the terms of *Carbon Capture and Usage (CCU)*, the captured CO_2 can also be reused for other processes [36,44,47].

Organisational success factor	**Description**
Definition and Implementation of a climate strategy	The *definition and implementation of a climate strategy* is an essential step towards achieving climate neutrality. With a climate strategy, the goal of climate neutrality is strategically anchored in the company and the corporate culture to raise the awareness of employees. Sub-goals are defined for example for certain departments or cost centres, and steps are specified to achieve them. [34,37,40,52,53]
Performing energy management	Since a large proportion of greenhouse gas emissions from companies are energy-related, effective *energy management* is necessary. Key aspects are, for example, the creation of transparency regarding energy. It is characterised by a continuous improvement process, such as PDCA described by the ISO 50001[54]. Systematic *energy management* supports the implementation of other success factors such as increasing energy efficiency. [8,35,39,40]
Incentivising emission reduction	*Incentivising emission reductions* may include the introduction of internal carbon pricing or the attribution of energy costs to departments according to their origin. To be effective, incentives for emission mitigation efforts should be communicated effectively. [55–57]
Supply chain engagement	A major part of the emissions arises from the production of goods in upstream processes outside the company's boundaries. They therefore belong to Scope 3 and can be influenced only indirectly by the company. However, companies can act on their supply chain and encourage suppliers to meet emission targets. Hence, *supply chain engagement* is considered a success factor [58].
Offsetting of remaining emissions	Despite the application of various success factors, companies often have emissions that cannot be avoided directly. In this case, the mandatory purchase of emission certificates or the voluntary *offsetting of remaining emissions* can be considered, which can at least reduce the climate-damaging effect of emissions [8,16,38–40,59]. However, the effectiveness of offsetting payments is discussed controversially [60].

Human success factor	**Description**
Commitment of the employees and management	The *commitment of the employees and management* is needed for the successful transformation towards climate neutrality [39]. It is crucial that they see climate change mitigation as a desirable goal [34]. Employee commitment impacts decisions towards sustainability at all levels, especially at the top management level [40,61].
Providing relevant competences	To enable employees to facilitate the development of the company towards climate neutrality and raise awareness for this strategic goal, *providing the relevant competences* or their external availability is necessary [40]. This is particularly the case in SMEs, where it is hardly possible for employees to specialise in the implementation of climate strategies [8,35,40].

4. Aligning shop floor management with the success factors for climate neutral production

After the detailed description of the literature results the relation of success factors for climate neutral production to SFM are analysed and discussed. The aim is to find out what element of SFM can actively support the success factors and to derive recommendations for further research. This is done by comparing the requirements for the implementation described in Table 1 with the elements of SFM. The findings are visualised in a matrix in Table 2.

In total 9 of the 16 success factors can be supported by SFM. These are described in detail in the following paragraphs. The remaining seven success factors cannot be incorporated into SFM. Those are outside of the sphere of influence of classic SFM. *Using green fuels* or *renewable energies* are strategic decisions made by the management or during a production planning process. The same holds for decisions to *implement aspects* *of circular economy* or *carbon capture and storage*. *Supply chain engagement* and the *compensation of remaining emissions* through the purchase of certificates is controlled by the purchasing department as they control where money is spent. Both are not subjects of a production process optimization method like SFM.

Table 2: Aligning SFM with the success factors of climate neutral production

		Success factors of climate neutral production															
		Technical									**Organisational**					**Human**	
		increasing energy efficiency	increasing energy flexibility	increasing material efficiency	reuse of waste heat	electrification of energy demands	using green fuels	using renewable energies	implementing aspects of circular economy	using carbon capture and storage/usage	implementing a climate strategy	performing energy management	incentivising emission reductions	supply chain engagement	compensation of remaining emissions	commitment of the employees	providing relevant competences to employees
Elements of SFM	**Performance Management** stabilization of processes, definition of goals and KPIs, control and increase of production output	●	●	●	●						●	●	●			●	●
	Problem Management systematic problem-solving process, continuous improvement	●		●	●												
	Lean Leadership Go and See, Gemba															●	●
	Glass Wall Management visual management, transparency, information dissemination										●	●	●			●	●
	Competence Management transfer of responsibility to employees, open communication, competence development										●					●	●

First, the technical success factors *increase of energy* and *material efficiency*, *energy flexibility* as well as *reuse of waste heat* are discussed. To *increase energy efficiency* one of the requirements is to identify potentials for energy reduction. Target values such as those used in performance management could be helpful for identifying potentials since target states are defined and deviations are determined. Deviations from the target state are dealt with in Problem Management. The *increase of energy flexibility* can be achieved by planning production according to energy availability. This is only possible when the processes are stable and therefore predictable. Through standardization in performance management this requirement can be achieved. For example, a standardized process with high energy demand is carried out at the time when the most renewable energy is available and therefore the energy costs are low. The *increase of material efficiency* is dependent on high quality in production processes. All forms of waste are to be eliminated which is in line with the goals of performance management as well as problem management. Another requirement to increase material efficiency is the low-waste design of products. It would be possible to use the proposal system of SFM for ideas to improve the design of components and products.

The first organizational success factor is the *implementation of a climate strategy*. Key elements of a climate strategy are the definition of goals for each department and awareness of every employee about their goals. Via Performance management these goals can be implemented on the shop floor. Through visual management and information dissemination (glass wall management) and an open communication (competence management) the strategy can be spread further and the chances of a successful change in corporate culture may increase. The second organizational success factor is to *perform energy management*. A requirement of an energy management system in line with ISO50001 is to achieve transparency of all energy processes. Once again KPIs in SFM can support this and elements of glass wall management like mini factories transparently display the energy processes. Furthermore, the ISO50001 recommends the use of PDCA (part of the SPSP) to manage energy related problems. For achieving an *incentivising of emission reductions* internal carbon pricing can be implemented. By making carbon emissions tangible in figures through internal pricing, a direct comparison with other areas is possible. The same strategy is applied in SFM with KPIs. A big part of internal carbon pricing is the effective communication to employees to make sure they understand why they are penalised for creating carbon emissions. Glass wall management in the form of mini factories could enhance transparency in that area.

To *provide relevant competences to employees* there needs to be a definition what the "relevant" competences for each employee are. Relevant competences are specified in SFM via the KPIs of each individual employee. The open and structured communication top-down as well as bottom-up as a main benefit of SFM can be a key feature for a successful transformation as well since it supports the requirement to make employees aware of the wastes around them. Furthermore, the goals of glass wall management could not only provide relevant information to develop competences but also support the *commitment of the employees* on the shop floor. On the other hand the commitment of management is shown transparently by Lean Leadership, where managers commit themselves to be active on the shop floor and support their employees.

5. Recommendations for further research

In summary, this paper analysed the suitability of SFM in the context of climate neutral production. After summarizing the success factors for the transformation to climate neutral production, the commonalities between SFM and the success factors were found. The paper identifies a need for further research and testing on two main topics. Topic one is that several success factors are dependent on KPIs to measure, control and improve resource efficiency of production processes. This requires new KPIs for climate neutral production which have played a subordinate role in classic SFM to date and have not yet been described in detail in this context. Following the recommendations by Diez et. al. [24], there is a need for a complete "Hoshin Kanri Tree" defining the goals and KPIs of climate neutral production on all hierarchical levels. A major obstacle here is the delimitation of the resource consumption caused by each individual workstation because of the difficulty to "break down" the savings from e.g., central supply systems that provide compressed air, heat, or cooling to several machines.

The second topic is the competences of employees and management in climate neutral production that need to be improved. To be able to use SFM for the transformation a full description of relevant competences for shop floor personnel as well as management is needed. These so-called competence tables can then be the basis for competence development through SFM.

Both steps form the basis to evaluate if the methods of SFM are suitable to support a climate neutrality strategy. The authors are planning to implement the new approach in a company experienced with classic SFM. The implementation will reveal which further modifications in the methods of SFM are needed to work with the new goal of transforming production towards climate neutrality.

Acknowledgements

This project (HA project no. 1012/21-14) is financed with funds of LOEWE – Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz, Förderlinie 3: KMU-Verbundvorhaben (State Offensive for the Development of Scientific and Economic Excellence).

References

[1] United Nations, 2015. Paris Agreement, Paris. http://unfccc.int/files/meetings/paris_nov_2015/application/pdf/paris_agreement_english_.pdf. Accessed 5 December 2021.

[2] United Nations, 2021. Glasgow Climate Pact, Glasgow. https://unfccc.int/sites/default/files/resource/cop26_auv_2f_cover_decision.pdf. Accessed 7 December 2021.

[3] Cherrafi, A., Elfezazi, S., Chiarini, A., Mokhlis, A., Benhida, K., 2016. The integration of lean manufacturing, Six Sigma and sustainability: A literature review and future research directions for developing a specific model. Journal of Cleaner Production 139, 828–846.

[4] Titmarsh, R., Assad, F., Harrison, R., 2020. Contributions of lean six sigma to sustainable manufacturing requirements: an Industry 4.0 perspective. Procedia CIRP 90, 589–593.

[5] Petry, M., Köhler, C., Zhang, H., 2020. Interaction analysis for dynamic sustainability assessment of manufacturing systems. Procedia CIRP 90, 477–482.

[6] Abualfaraa, W., Salonitis, K., Al-Ashaab, A., Ala'raj, M., 2020. Lean-Green Manufacturing Practices and Their Link with Sustainability: A Critical Review. Sustainability 12 (3), 981.

[7] Honegger, M., Schäfer, S., Poralla, P., Michaelowa, A., Perspectives Climate Research gGmbH, 2020. dena-Analyse: Klimaneutralität – ein Konzept mit weitreichenden Implikationen, Freiburg i. B. https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2020/dena_BR_Analyse-Klimaneutralita__t_WEB.pdf. Accessed 1 December 2021.

[8] 2021. dena-Leitstudie Aufbruch Klimaneutralität, Berlin. https://www.dena.de/fileadmin/dena/Publikationen/PDFs/2021/Abschlussbericht_dena-Leitstudie_Aufbruch_Klimaneutralitaet.pdf. Accessed 4 December 2021.

[9] Leyendecker, B., Pötters, P., 2018. Shopfloor Management: Führung am Ort des Geschehens. Hanser, München, 125 pp.

[10] European Commission, 2019. The European Green Deal, Brussels. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2019:640:FIN. Accessed 1 December 2021.

[11] European Commission, 2021. 'Fit for 55': delivering the EU's 2030 Climate Target on the way to climate neutrality, Brussels. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021DC0550. Accessed 1 December 2021.

[12] Millot, A., Krook-Riekkola, A., Maïzi, N., 2020. Guiding the future energy transition to net-zero emissions: Lessons from exploring the differences between France and Sweden. Energy Policy 139, 111358.

[13] Buettner, S.M., König, W., 2021. Looking behind decarbonisation - What pressure points trigger action?, in: eceee Summer Study Proceedings, Stockholm, pp. 345–354.

[14] World Resources Institute, 2015. The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard, Washington. https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf. Accessed 1 December 2021.

[15] World Resources Institute, 2011. Corporate Value Chain (Scope 3) Accounting and Reporting Standard: Supplement to the GHG Protocol Corporate Accounting and Reporting Standard, Washington. https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf. Accessed 1 December 2021.

[16] Prognos, Öko-Institut, Wuppertal-Institut, 2020. Klimaneutrales Deutschland: Klimaneutrales DeutschlandIn drei Schritten zu null Treibhausgasen bis 2050 über ein Zwischenziel von -65 % im Jahr 2030 als Teil des EU-Green-Deals. https://static.agora-energiewende.de/fileadmin/Projekte/2020/2020_10_KNDE/A-EW_195_KNDE_WEB.pdf. Accessed 4 December 2021.

[17] International Energy Agency, 2021. Net Zero by 2050: A Roadmap for the Global Energy Sector. https://iea.blob.core.windows.net/assets/deebef5d-0c34-4539-9d0c-10b13d840027/NetZeroby2050-ARoadmapfortheGlobalEnergySector_CORR.pdf. Accessed 27 January 2022.

[18] Agora Energiewende und Wuppertal Institut (2019. Klimaneutrale Industrie: Schlüsseltechnologien und Politikoptionen für Stahl, Chemie und Zement, Berlin. https://static.agora-energiewende.de/fileadmin/Projekte/2018/Dekarbonisierung_Industrie/164_A-EW_Klimaneutrale-Industrie_Studie_WEB.pdf. Accessed 28 January 2022.

[19] Metternich, J., Schebek L., Anderl R., 2021. Fit Für Die Zukunft: Ressourceneffizienz in Produktionsprozessen. TU Darmstadt. https://www.arepron.com/index.php/publikationen.

[20] Suzaki, K., 1993. The new shop floor management: Empowering people for continuous improvement. The Free Press, New York.

[21] Peters, R., 2009. Shopfloor Management: Führen am Ort der Wertschöpfung. LOG_X, Ludwigsburg, 160 pp.

[22] Hertle, C., Tisch, M., Metternich, J., Abele, E., 2017. Das Darmstädter Shopfloor Management-Modell. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 112 (3), 118–121.

[23] Hertle, C., Tisch, M., Kläs, H., Metternich, J., Abele, E., 2016. Recording Shop Floor Management Competencies – A Guideline for a Systematic Competency Gap Analysis. Procedia CIRP 57, 625–630.

[24] Diez, J.V., Ordieres-Mere, J., Nuber, G., 2015. The HOSHIN KANRI TREE. Cross-plant Lean Shopfloor Management. Procedia CIRP 32, 150–155.

[25] Meißner, A., Grunert, F., Metternich, J., 2020. Digital shop floor management: A target state. Procedia CIRP 93, 311–315.

[26] Blöchl, S.J., Michalicki, M., Schneider, M., 2017. Simulation Game for Lean Leadership – Shopfloor Management Combined with Accounting for Lean. Procedia Manufacturing 9, 97–105.

[27] Meissner, A., Müller, M., Hermann, A., Metternich, J., 2018. Digitalization as a catalyst for lean production: A learning factory approach for digital shop floor management. Procedia Manufacturing 23, 81–86.

[28] Hertle, C., et al., 2017. Innovative Approaches for Technical, Methodological, and Socio-communicative Competency Development in Production Areas. Procedia Manufacturing 9, 299–306.

[29] Hertle, C., Siedelhofer, C., Metternich, J., Abele, E., 2015. The next generation shop floor management – how to continuously develop competencies in manufacturing environments. Universitäts- und Landesbibliothek Darmstadt, Darmstadt, Online-Ressource.

[30] Makhija, A., Wickramasinghe, C., Tiwari, M., 2021. 5 - Visual management, in: Jana, P., Tiwari, M. (Eds.), Lean Tools in Apparel Manufacturing : The Textile Institute Book Series. Woodhead Publishing, pp. 131–208.

[31] Müller, M., Alexandi, E., Metternich, J., 2021. Digital shop floor management enhanced by natural language processing. Procedia CIRP 96, 21–26.

[32] Gaitzsch, T., Ziegler, V., 2010. Shop Floor Empowerment – KVP-Implementierung in Schichtteams, in: Moscho, A., Richter, A. (Eds.), Inhouse-Consulting in Deutschland: Markt, Strukturen, Strategien. Gabler, Wiesbaden, pp. 169–190.

[33] Strohm, O., 1997. Unternehmen arbeitspsychologisch bewerten: Ein Mehr-Ebenen-Ansatz unter besonderer Berücksichtigung von Mensch, Technik und Organisation. vdf Hochschulverlag AG an der ETH Zürich, Zürich, 511 pp.

[34] Carlsson Kanyama, A., Carlsson Kanyama, K., Wester, M., Snickare, L., Söderberg, I.-L., 2018. Climate change mitigation efforts among transportation and manufacturing companies: The current state of efforts in Sweden according to available documentation. Journal of Cleaner Production 196, 588–593.

[35] Sovacool, B.K., et al., 2021. Decarbonizing the food and beverages industry: A critical and systematic review of developments, sociotechnical systems and policy options. Renewable and Sustainable Energy Reviews 143, 110856.

[36] Zou, C., et al., 2021. Connotation, innovation and vision of “carbon neutrality”. Natural Gas Industry B 8 (5), 523–537.

[37] van Sluisveld, M.A., de Boer, H.S., Daioglou, V., Hof, A.F., van Vuuren, D.P., 2021. A race to zero - Assessing the position of heavy industry in a global net-zero CO2 emissions context. Energy and Climate Change 2, 100051.

[38] Bataille, C., J. Nilsson, L., Jotzo, F., 2021. Industry in a net-zero emissions world: New mitigation pathways, new supply chains, modelling needs and policy implications. Energy and Climate Change 2, 100059.

[39] Rissman, J., et al., 2020. Technologies and policies to decarbonize global industry: Review and assessment of mitigation drivers through 2070. Applied Energy 266, 114848.

[40] Choi, J.-K., Schuessler, R., Ising, M., Kelley, D., Kissock, K., 2018. A Pathway Towards Sustainable Manufacturing for Mid-size Manufacturers. Procedia CIRP 69, 230–235.

[41] Sauer, A., Abele, E., Buhl, H.U. (Eds.), 2019. Energieflexibilität in der deutschen Industrie: Ergebnisse aus dem Kopernikus-Projekt - Synchronisierte und energieadaptive Produktionstechnik zur flexiblen Ausrichtung von Industrieprozessen auf eine fluktuierende Energieversorgung (SynErgie). Fraunhofer Verlag, Stuttgart, 728 pp.

[42] Reinhart, G., Reinhardt, S., Graßl, M., 2012. Energieflexible Produktionssysteme: Einführungen zur Bewertung der Energieeffizienz von Produktionssystemen. wt Werkstattstechnik online (9), 622–628.

[43] Thiel, G.P., Stark, A.K., 2021. To decarbonize industry, we must decarbonize heat. Joule 5 (3), 531–550.

[44] Capros, P., et al., 2019. Energy-system modelling of the EU strategy towards climate-neutrality. Energy Policy 134, 110960.

[45] Imdahl, C., et al., 2021. Potentials of Hydrogen Technologies for Sustainable Factory Systems. Procedia CIRP 98, 583–588.

[46] Naegler, T., Sutardhio, C., Weidlich, A., Pregger, T., 2021. Exploring long-term strategies for the german energy transition - A review of multi-Sector energy scenarios. Renewable and Sustainable Energy Transition 1, 100010.

[47] Ipsakis, D., et al., 2021. Techno-economic assessment of industrially-captured CO2 upgrade to synthetic natural gas by means of renewable hydrogen. Renewable Energy 179, 1884–1896.

[48] Wang, F., et al., 2021. Technologies and perspectives for achieving carbon neutrality. The Innovation 2 (4), 100180.

[49] Gerres, T., Chaves Ávila, J.P., Linares Llamas, P., Gómez San Román, T., 2019. A review of cross-sector decarbonisation potentials in the European energy intensive industry. Journal of Cleaner Production 210, 585–601.

[50] Turner, K., Race, J., Alabi, O., Katris, A., Swales, J.K., 2021. Policy options for funding carbon capture in regional industrial clusters: What are the impacts and trade-offs involved in compensating industry competitiveness loss? Ecological Economics 184, 106978.

[51] Canal Vieira, L., Longo, M., Mura, M., 2021. Are the European manufacturing and energy sectors on track for achieving net-zero emissions in 2050? An empirical analysis. Energy Policy 156, 112464.

[52] Lee, H., 2021. Is carbon neutrality feasible for Korean manufacturing firms?: The CO2 emissions performance of the Metafrontier Malmquist–Luenberger index. Journal of Environmental Management 297, 113235.

[53] Zameer, H., Wang, Y., Vasbieva, D.G., Abbas, Q., 2021. Exploring a pathway to carbon neutrality via reinforcing environmental performance through green process innovation, environmental orientation and green competitive advantage. Journal of Environmental Management 296, 113383.

[54] Energy management systems - Requirements with guidance for use (ISO 50001:2018).

[55] Harpankar, K., 2019. Internal carbon pricing: rationale, promise and limitations. Carbon Management 10 (2), 219–225.

[56] Bento, N., Gianfrate, G., 2020. Determinants of internal carbon pricing. Energy Policy 143, 111499.

[57] Gorbach, O.G., Kost, C., Pickett, C., 2022. Review of internal carbon pricing and the development of a decision process for the identification of promising Internal Pricing Methods for an Organisation. Renewable and Sustainable Energy Reviews 154, 111745.

[58] Mahapatra, S.K., Schoenherr, T., Jayaram, J., 2021. An assessment of factors contributing to firms' carbon footprint reduction efforts. International Journal of Production Economics 235, 108073.

[59] Stede, J., Pauliuk, S., Hardadi, G., Neuhoff, K., 2021. Carbon pricing of basic materials: Incentives and risks for the value chain and consumers. Ecological Economics 189, 107168.

[60] Barron, A.R., Domeshek, M., Metz, L.E., Draucker, L.C., Strong, A.L., 2021. Carbon neutrality should not be the end goal: Lessons for institutional climate action from U.S. higher education. One Earth 4 (9), 1248–1258.

[61] Kohn, K., Seyfried, S., Weyand, A., Weigold, M., 2021. Energieeffizienz als ein Ausdruck der Nachhaltigkeit produzierender Betriebe. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (1-2), 34–38.

Biography

Sebastian Bardy, M. Sc. (*1991) has been PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2019. His research interests include the introduction and further development of resource-efficiency in digital shop floor management.

Stefan Seyfried, M. Sc. M. Sc. (*1990) has been PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2018. His research interests include energy efficiency and climate strategies of industrial enterprises and SMEs.

Prof. Dr.-Ing. Joachim Metternich (*1968) has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2012. He is spokesman of the medium-sized businesses digital competence centre Darmstadt; he has been president of the International Association of Learning Factories (IALF) from 2015 to 2021 and he is a member of the scientific society for production technology (WGP).

Prof. Dr.-Ing. Matthias Weigold (*1977) has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at Darmstadt University of Technology since 2019. As director he is responsible for the research fields on Manufacturing Technology as well as Energy Technologies and Applications in Production.

3rd Conference on Production Systems and Logistics

Development Of A Method For Decision Support On Participation In Capacity Sharing For Manufacturing SMEs

Maren Müller[1], Fabian Schüler[2] , Malte Stonis[1] , Peter Nyhuis[1]
[1] *IPH – Institut für Integrierte Produktion Hannover gGmbH, Hannover, Germany*
[2] *IPRI – International Performance Research Institute gGmbH, Stuttgart, Germany*

Abstract

A volatile, non-transparent market environment leads to fluctuations in the load on production capacities in the manufacturing sector, which are reflected within production in over- or underutilization of machines and persons. Small and midsized enterprises (SMEs) are expecting increasing volatility, which is accompanied by an increase in the frequency of market and economic cycles. For SMEs it is difficult to handle these fluctuations. Capacity sharing platforms can be a solution for this challenge. Platforms are available in different forms. Currently, companies are not often using this possibility, because of prevailing scepticism in different fields. Therefore, a methodology will be developed to provide a decision support for or against platform usage. Additionally, the platform type choice will be supported, and the changes of logistic and economic indicators will be considered. With this information, companies can make a qualitative decision, and the existing inhibitions can be alleviated.

Keywords

Capacity sharing; decision support; supplier and consumer view; logistic and economic indicators

1. Introduction

The digitization of the economy across industries is not only leading to the emergence of innovative products and services, but also to a transformation of existing market logics [1]. Digital platforms as a growing innovation driver of digital transformation exists [2]. By linking the sharing economy in the form of capacity sharing with digital platforms, companies can offer or purchase their free or required capacities on an inter-company basis. This potential flexibility is currently getting more and more attention. SMEs are expecting increasing volatility, which goes hand in hand with an increase in the frequency of market and economic cycles [3]. Especially for SMEs, these fluctuations are difficult to map internally, so capacity sharing is seen as a solution to compensate these fluctuations. The decision of a beneficial of using capacity sharing for an individual company and which type of platform is suitable, depends on various factors. Therefore, a methodology has been developed to help companies to make this decision. It also provides a guidance on how economic and logistic indicators would likely change by using capacity sharing. Companies can make a high-quality decision based on this provided information in a user-friendly application including a guidance [4]. *

DOI: https://doi.org/10.15488/12197

2. State of the art

Many SMEs are currently affected by increasing workload fluctuations. The state of the art will first present, how the current way looks like to challenge these fluctuations, followed by a presentation of the current use of capacity sharing in companies.

2.1 Workload fluctuations in SMEs and current compensation

Most recently, more than a half of all German SMEs were affected by sales losses averaging 53% compared with expected sales in March 2020 [5]. The reasons for this are for example the measures to combat the pandemic or the declines in demand for orders [5]. A look at the change in production in the manufacturing sector in Germany compared with the previous month from January 2020 to 2021 shows significant fluctuations in orders and thus capacity utilization [6]. A study by GREAN GmbH in the issue of "Production after the pandemic" shows, that there is a high level of capacity utilization during the pandemic and most companies also expect a rapid recovery [7]. The high level of capacity utilization is due to an artificial reduction in available capacity, e.g. through short-time working, and to existing orders that were placed before the outbreak of the pandemic [8].

Independently of pandemics, companies are confronted with fluctuations in capacity utilization. The reasons can be found in a volatile market and the globalization [9]. Furthermore, fluctuations in capacity utilization can be influenced by the economic situation, e.g., during the financial crisis in 2009 [10]. Additionally, political conditions or natural disasters can lead to material shortages [11]. However, the sector itself also entails a certain degree of fluctuation in capacity utilization, for example due to seasonal products [12].

To counter fluctuations in capacity utilization, companies can make use of various flexibility instruments. These can include working time accounts, short-time work, temporary work, extending shift work or increasing weekly working hours [13]. Furthermore, permanent machine availability can be ensured by an in-house technical support. These measures counteract the consequences mainly by adjusting the available internal personnel capacities [13]. However, if these are fully exhausted or if no plant and machinery is available, or the disadvantages for the persons are to be abolished, the use of capacity sharing can be beneficial. This applies to the situation where free capacities are available in the own company and can therefore be offered as well as to an overload and the subsequent external assignment of orders.

2.2 Capacity sharing for SMEs in the manufacturing sector

Fluctuations in capacity utilization are increasingly challenging manufacturing companies. A digital platform makes it possible to exchange production capacities to compensate fluctuations in capacity utilization. It is possible to participate in a platform as a supplier, a consumer or both, depending on the current situation in the companies.

In the private environment, the sharing economy is becoming increasingly important. Opportunities for car sharing and platforms for renting accommodation are widely used [14]. In the industry, there are opportunities for capacity sharing, which are little used. Reasons for the restrained use can be found, for example, in the competitive situation of the companies. The decision to outsource production steps or to use a platform is often equated with the disclosure of own production ideas [15]. But it is also difficult to assess the uncertainties about the course and cost of such an order, the quality of the externally produced product or the effects on the company's own production processes.

Currently, there are several active platforms that offer different manufacturing processes such as milling, drilling, 3D printing or CAD design for different batch sizes. The possible customer groups are very diverse and cross-industry. Since the technical possibilities are available in a variety of different solutions, the companies must now be picked up, comprehensively informed, and supported.

The need for companies to be adaptable is seen and described in the literature. However, the flexibility instruments concern internal personnel, which creates only internal solutions. Due to the large number of platform providers, an external and cross-company solution could be provided. Nevertheless, this is mostly not used by SMEs because of uncertainties whether capacity sharing makes sense for the companies and which platform should be used. Furthermore, the effects on their own production are unknown. Currently, there is no scientific support for companies to decide whether the use of capacity sharing is individually useful, which capacity sharing platform is suitable and what are the effects on their production. This paper aims to fill the research gap by developing a method for decision support on participation in capacity sharing for manufacturing SMEs.

3. Decision support for the use of capacity sharing

The structure of the developed decision support is shown in figure 1.

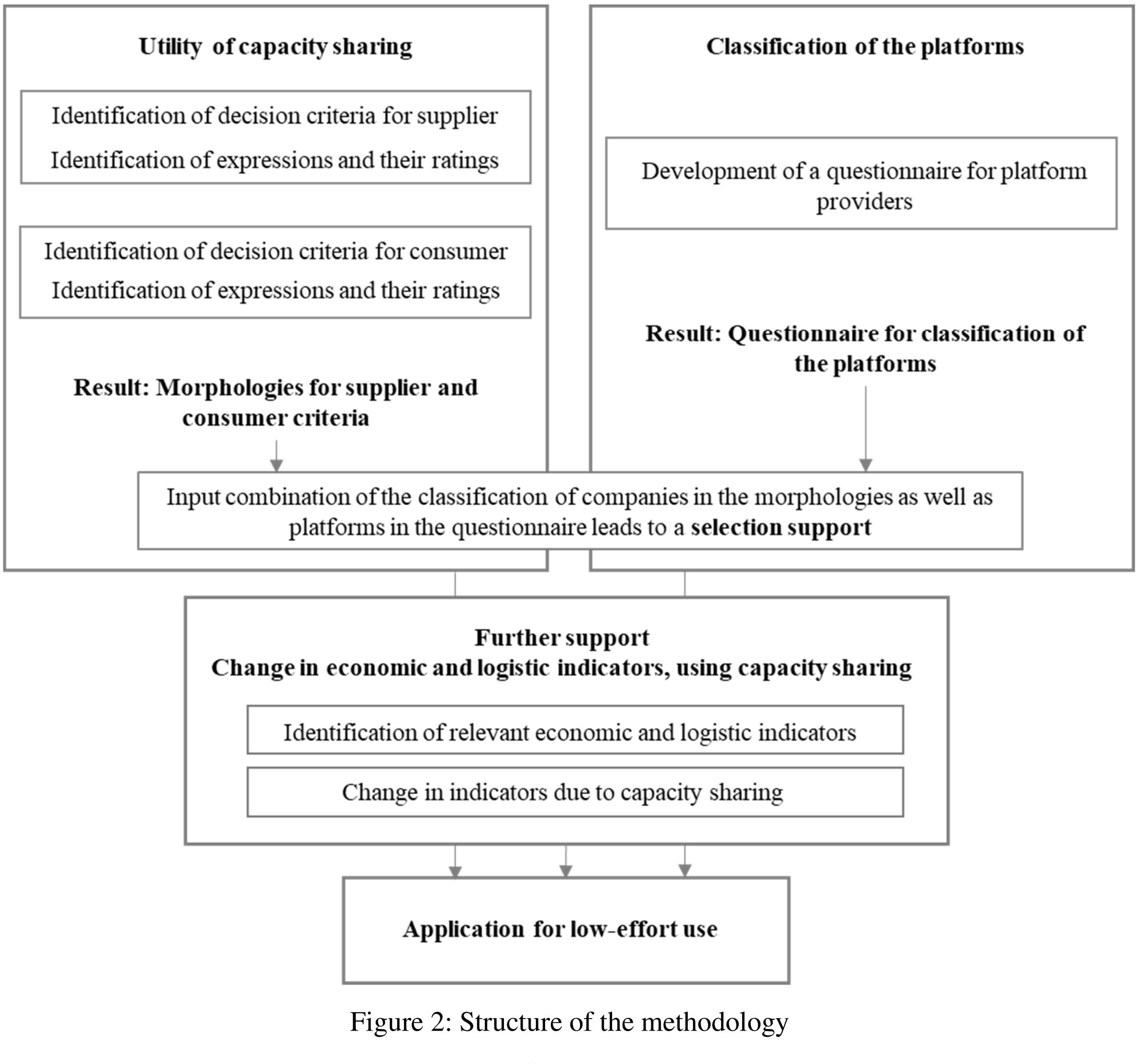


Figure 2: Structure of the methodology

To provide comprehensive decision support, the first step is to develop a methodology that can be used to evaluate, whether it is beneficial to participate in capacity sharing as a supplier/consumer. For this, decision criteria must be identified as well as their expressions and ratings. As a result, morphologies arise for supplier and consumer criteria. In a second step, support is provided to decide which platform is suitable for a company and its requirements. For this, a questionnaire must be developed for the platform providers. As a result a classification can be create for each platform. These steps can be combined to create a platform selection support. Depending on the placement of the companies to the expressions of the decision criteria as well as the questionnaire response of the platform providers, a suitable platform can be suggested. Additionally, the change in economic and logistic indicators, arise by using capacity sharing, will be presented. Therefore, relevant economic and logistic indicators were identified and the change due to capacity sharing will be presented. In the last step, the selection support and further information will be integrated and presented in an application for a low-effort use.

3.1 Utility of capacity sharing

To evaluate participation in a capacity sharing platform, decision criteria must be defined as well as their expressions. These will be presented in a morphology. Thereby, a morphological box is created for the supplier and consumer side. Already here, attention was paid to the following task, which deals with the assignment of a suitable platform and considers criteria in this topic too.

3.1.1 Identification of decision criteria and expressions for supplier

Table 2 shows the final supplier morphology, which was developed via iterative interviews with experts[1]. The forms of expression of the individual dimensions already give an indication of which characteristics stands for (column 3) and which stands against (column 1) participating in capacity sharing platforms. The middle column shows indicators which must be considered, when selecting a capacity sharing platform, but which do not give a clear indication for or against the use, but offers properties to select a suitable platform.

The **fluctuations in demand,** postulated in this paper, are confirmed by the experts, and therefore enable temporary capacity sharing as an attractive application scenario. Also, the **willingness to increase the machine utilization** should be given. As far as the **availability of utilization data is concerned**, most of the companies do not currently have any (usable) data regarding to the utilization of individual machines. For this reason, automated a matching based on real utilization data can only be implemented in a few scenarios or by retrofitting the machines accordingly [16]. Due to the sometimes-large effort required for retrofitting and various critical security aspects about data provision, an alternative solution approach to capacity matching is the active maintenance of integrated production planning, in which corresponding slots for individual machines can be enabled or blocked.

An evaluation of the **availability of various manufacturing processes** reveals major differences, which must be considered. Discussions with the experts identified the manufacturing processes, which are currently requested via platforms or offer potential for use.

The **quantity and variety** of the raw material in inventory represents another decision dimension that must be queried in an automated matching process. Ideally, the available quantities of raw material should be read directly from a producer's system during the matching process. If an order is placed by means of individual quotations from the production partners, the check can be carried out as a part of the quotation preparation process. However, since it is not known in advance, which company has which materials and in which extent, the overall effort requires to prepare the offer increases and the efficiency of the platform suffers. It is recommended for the production partners to maintain a virtual material inventory, which can be considered during the matching.

The **infrastructure** should give the possibility to add further orders for using capacity sharing platforms. The general **flexibility** of the organization as well as the general **willingness** of the companies to support an additional platform are queried, as these characteristics represent a basic condition for the use of a capacity sharing platform.

---

[1] The team of experts consists of employees of manufacturing companies (5), consulting companies (3) and research institutes / associations (4) as well as platform providers (3). This team of experts is also meant in the further course of this paper.

Table 1: Supplier morphology

Characteristic		Eexpressions		
		1	2	3
Do fluctuations in demand frequently lead to temporary underutilization of machines?			no	yes
Do you have machines, whose utilization you would like to increase in general?			no	yes
Is your company's production capacity utilization data available?		unknown	unknown, not standardized	standardized, digital available
Utilization of individual production areas	Additive manufacturing (ceramic-based)	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Additive manufacturing (metal-based)	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Additive manufacturing (polymer-based)	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Sheet metal processing (laser cutting, bending, surface treatm.)	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	CNC-turning	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	CNC-milling	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Plastic molding (injection molding, extrusion)	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Tube processing / tube bending	high utilization/ not relevant	overload during order peaks	low to medium Utilization
	Welding	high utilization/ not relevant	overload during order peaks	low to medium Utilization
Quantity and variety of available stocked raw material for the specified manufacturing areas.		low		high
Additional orders can be added (infrastructure).		no		yes
How high do you rate the flexibility of your production planning and control system?		low		high
Personnel utilization	Construction	high		low
	Manufacturing	high		low
	Assembly	high		low
Supporting Areas	Is there a possibility in the WMS to consider additional orders?	no		yes
	How high is the effort to integrate additional orders in the material flow?	high		low
	How high is the effort to consider additional orders in the goods issue?	high		low
High number of tenders with low chance of success or few tenders with high chance of success?		many offers great competition		few, individual offers
Would you like much transparency and comparability about the processes and services of the platform?		no		yes
Knowledge/skills		Certificates not available		Certificates available

The **personnel utilization** in production represents another important decision criterion for capacity sharing. If there is a high utilization of personnel despite low utilization of individual machines, the orders must be selected in such a way that the pure production times, in which the machine works autonomously, dominate, in relation to the work preparation (setup, clamping, reclaiming, etc.). This can be achieved by particularly high machining times per part or large quantities of individual jobs. In the future, capacity sharing platforms should take this aspect into account.

In the **supporting areas**, serious differences also become apparent between companies, that manufacture their own products and contract manufacturers. While the latter have hardly any restrictions in the supporting areas and the processes are optimized for the short-term processing of external orders, the material flow-oriented organization of manufacturing companies leads to serious restrictions about capacity sharing in the supporting processes. Here, new processes often must be developed to enable the integration of new special orders into the material flow, warehouse management (WMS) and goods issue processes.
Also important is the preferred scenario about the **number of offers and the competition** as well as the availability of **transparency and comparability**, which effects the decision for one special platform.
About the **necessary knowledge and skills,** there is also a widespread among the manufacturing partners surveyed. While all potential manufacturing partners have at least one external certification (mostly ISO 9001), the companies have numerous other certifications that must be considered in the matching process. Relevant certifications exist, for example, for the automotive industry (IATF 16949), aerospace (EN 9001) or medical technology (ISO 13485). Due to liability issues across the supply chain, the central task of the platform here is to evaluate the production partners' certificates to fulfil these requirements. In table 1, the differentiation between no certificates and certificates available is shown in a simplified form. In the resulting matching, the certificates must be considered in more detail.

3.1.2 Identification of decision criteria and expressions for consumer

Table 2 presents the relevant decision criteria from the consumer perspective.

A central distinction is already provided by the question, whether the outsourcing part is a mission-critical process or a **core competence** of the company. In these cases, reservations about capacity sharing platforms are significantly greater than in the case of the production of spare parts, prototypes, or custom-made products from special mechanical engineering. In the case of the former, liability issues must be clarified (product liability, intellectual property), which is why close cooperation must be established between company's purchasing department and the capacity sharing platform.
In the indirect areas, capacity sharing platforms offer a major advantage when a company's **purchasing department** is working utilized. Through the platform, more **purchasing autonomy** can be assigned to individual departments (testing, prototype, construction), which relieves the indirect areas to some extent. Also, the experience of companies in the field of assigning **external partner** for an extern production could be a relevant indicator.
The next characteristic is about the **design data**. The digital availability, the design data quality and the format are important for the usage of a capacity sharing platform. Since all capacity sharing platforms are fully digitized, the provision of design data in step format (Standard for the Exchange of Product model data) is usually recommended or, in some cases, assumed as a minimum requirement.
The reduction of the **geographical distance** of the exchange partners is also important for many German companies. In addition to reducing costs, the **environmental protection** and support of global sustainability goals [17].
Another important decision dimension is the **characteristics of product properties** of the orders to be placed. Important is the question, whether standard or special material must be used. In addition, it is decisive, whether the manufacturing process is exclusively for a single part or whether additional assembly steps are necessary for the manufacturing of assemblies or entire products. Assembly and the associated material procurement increase the complexity of matching enormously. Furthermore, it is also decisive whether semi-finished products must be provided for production or whether the initial process step is mediated. This is important because only a few platforms have a standardized process for the provision of input material.
Important for the decision for one special platform type is the preferring **price possibilities** (instant price and comparison offers).

Table 2: Consumer morphology

Characteristic		Expressions		
		1	2	3
Core competence / success-critical process		yes		no
Utilization personnel purchasing		low		high
Would you like to increase the purchasing autonomy / flexibility of individual departments?		no		yes
Do you often need external partners to compensate for order peaks?		no		yes
Design data	Digital availability	no		yes
	Design data quality	low		high
	The data is transmitted as a step file. Should orders also be placed by means of a technical drawing (pdf)?	Partial orders via pdf		no
Importance of the environmental protection and CO^{2} balance of the production partner.			not important	important
Geographical distance of production partners.			important	not important
Characteristics of potential contracts to be awarded	Order material	special material		standard material
	Produkt type	produkt	module	component
	Is special input material required?	Provision Semifinished product		no starting material/initial process step
	Heterogeneity of orders	low		high
	Average contract value of contracts to be awarded	> 10.000€	2.000 - 9.999 €	< 2.000 €
Do you prefer a binding immediate price or do you require several comparative offers?		instant price		comparison offers
Data protection/ Data security	Desired server location of the platform	Germany	Europe	worldwide
	Certification according to ISO/IEC 27001 or 27002	required	desirable	not relevant
Flexibility		low		high
Legal framework conditions (e.g. standards and certificates to be met)		available		not available

3.1.3 Usage to identify the utility of capacity sharing

The characteristic **data protection and data security** is also important, especially for the decision for one platform, as well as the **flexibility** itself.

Finally, as with any other types of outsourcing, the **legal framework conditions** must be checked to determine the extent, to which a specific component can be earmarked for outsourcing. Most platforms have a standardized non-disclosure agreement (NDA), which can be viewed in advance on the homepage.

In the next step, an individual company can select the expressions of the criteria. As a first result, the user receives a supplier and consumer score for the suitability in percent. The classification of the resulting recommendation is shown for an example in Figure 1. In discussion with experts and the first company results, by using the decision support, the borderline between the suitability and a necessary further examination was set at 30 percent. This borderline is not fixed, it represents only an orientation. In further validation steps, this borderline must be analysed in more detail.

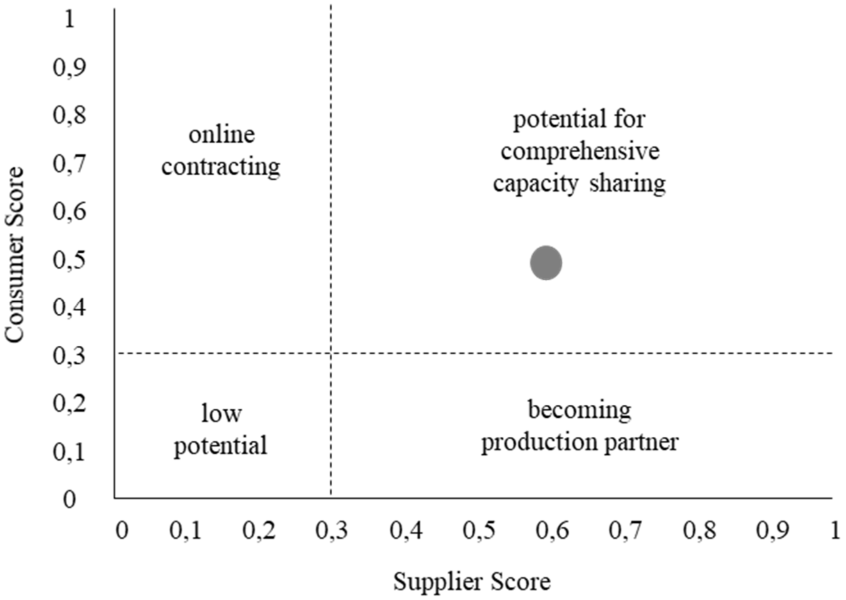


Figure 2: Illustration of the suitability for capacity sharing

In the present example, capacity sharing offers a great potential for the company, whereby the company can act as a supply and a consumer.

3.2 Classifications of the platforms

For the identification of the decision criteria for a platform type, the previously identified criteria were used as a basis. The two task areas were worked through in parallel, so that the companies do not have to provide any additional information. In consultation with platform providers, the differentiation possibilities of the platforms were worked out and integrated into the morphologies. The platform providers get a questionnaire, where they can provide information about their platforms. The following information should be provided by the platform providers:

- Supporting manufacturing types of the platform: additive manufacturing (ceramic-based, metal-based, and polymer-based), sheet metal working, CNC turning, CNC milling, plastics processing, tube processing, welding
- Supplier and consumer role possibilities
- Manufacturing of assemblies and/or individual products and/or individual work steps
- Preferred order volume
- Instant quotes / tender platform
- Transparency (e.g. FAQ)
- Orders via PDF possibility
- CO^2 neutrality
- Server location / data security (choose between: Germany, Europe, worldwide)
- ISO /IEC 27001 or 27002 - certifications (choice between: mandatory, desirable, not required)
- Hypertext Transfer, Protocol Secure (HTTPS) available

Based on the classification of the companies into the morphological boxes presented above and the answers to the questionnaire of the platform providers, an individual suitability to the platforms (in percent) can be determined. An actual overview of the platform providers was made available for this purpose.

3.3 Change in economic and logistic indicators, using capacity sharing

For the final evaluation, whether participation in a capacity sharing platform is beneficial for a company, the companies should know, how their economic and logistic indicators are likely to change. For this purpose, the relevant indicators were identified in a first step and ranked according to the strength of a possible change. In a second step, the changes for supplier and consumer were analysed, depending on the company scenario or initial situation. For this purpose, standard scenarios were developed with experts, for each of them the anticipated changes can be worked out.

3.3.1 Identification of relevant economic and logistic indicators

To identify the relevant economic and logistic indicators, all economic and logistic indicators were first provided. A list of relevant indicators for capacity sharing was then compiled with the involvement of experts (Table 3).

Table 3: Relevant economic and logistic indicators

economic indicators	logistic indicators
transport costs	delivery time
planning and control costs	downtime
storage costs	throughput time
production costs	delivery reliability
igle capacity costs	machine utilization
quality costs	personnel utilization
	inventory
	batch size
	setup time

The delivery time indicates the period of time that elapses from the placing of the order to its fulfilment [18]. Downtimes include technical malfunctions and other downtimes that were not scheduled [19]. **Throughput time** indicates the amount of time required from the start of production to completion. This includes idle time, setup time and processing time, as well as transport time, which reflects the distance between two workplaces [20, 21]. The adherence to promised delivery dates is described by the key figure **delivery reliability** [22]. **Machine utilization** indicates the ratio of actual machine working time to total available working time. The **personnel utilization** expresses the utilized personnel capacity in relation to the available working time [19]. **Inventories** give rise on the one hand to capital commitment costs, since goods held in inventory have to be financed, and on the other hand to storage costs [22]. The **batch size** is the quantity of products or parts that can be produced directly one after another without interrupting production [23]. **Setup time** is the time required to prepare a machine for the production of another variant [19].

Logistics costs are caused by the provision of a logistic service. In this paper, it includes the three superordinate service areas: **Transportation, storage and production planning and control (PPC)**, which are considered individually due to the different influences and anticipated changes. Transportation costs are classified as the costs incurred by the spatial change of goods. Storage costs are represented by warehousing, storage, and retrieval as well as the provision of storage space. PPC costs are incurred due to the rescheduling effort involved in taking on placing external orders [22, 18]. **Production costs** describe direct costs due to processing as well as maintenance, workshop and production performance costs [19]. **Idle capacity costs** arise from the non-utilization of existing capacity and therefore also reflect a certain degree of underemployment [24]. **Quality costs** arise from quality assurance or the restoration of the required quality through rework [19].

After identifying the relevant indicators, the next step is to assess the potential change itself. Some indicators are likely to change more than others, when capacity sharing is used. Table 4 shows the classification into high, medium, and low. This classification was made by discussions with experts.

These indicators can lead to both a positive and a negative change. It depends on the scenario or the present situation of a company as well as the role (supplier and/or consumer).

Table 4: Classification of indicators according to the strength of a possible change

high	medium	low
planning and control costs	production costs	storage costs
igle capacity costs	delivery reliabilty	inventory
machine utilization	batch size	
personnel utilization	setup time	
delivery time	quality costs	
downtime		
throughput time		

3.3.2 Change in indicators due to capacity sharing

To be able to finally evaluate the participation in a capacity sharing platform, the resulting changes in the economic and logistic indicators should be known. For the evaluation of these, a spreadsheet was developed, where the respective change for supplier and consumer, depending on the scenario or initial situation, is shown. For this purpose, expert interviews were conducted with users and capacity sharing platform holders. In total, the experts identified 10 different scenarios. Furthermore, a simulation model was built to show the changes in a simulative way to confirm the previously assumed changes. To illustrate the changes in this paper, a standard scenario is considered below that describes a manufacturing company that can act as a supplier and a consumer. The company uses the platform to compensate their seasonal fluctuations in orders. In addition, internal company influences are described that can affect the strength of the change.

The **transport costs** for the supplier or consumer increase due to the distance of the partner company and the additional transport effort. The product itself is a factor in the level of these costs. The bulkiness, volume and weight of a product are decisive factors. Another factor is the transport infrastructure between the partner companies.

The **PPC** costs increase in the supplier process and the consumer process, since the additional orders must be adjusted in the PPC. One factor for the level of this influence is the presence of system support, e.g., ERP or ME systems [25].

The behaviour of **storage costs** itself for the supplier and consumer does not normally differ from the original situation, but the capital commitment cost can be optimized.

The **production costs** for the supplier for additional orders are normally unchanged from the normal production costs of an own order, furthermore the fixed costs recovery can be optimized. Factors according to the strength of the change results is the number of additional orders as well as the deviations from the own products. The costs for the production as a consumer, in comparison with the own production, will be higher.

The suppliers' **idle capacity costs** can be drastically reduced. For a consumer, the idle capacity costs are already very low, when an outsourcing is used.

The **quality costs** for the suppliers normally remain unchanged. Only an additional necessary control could increase the costs. Increasing quality costs are likely to be incurred for the consumer, if the products are not shipped directly to their customer. After receipt of the products, they are checked in more detail before being further processed or shipped.

Participation can potentially have a negative impact on the **delivery time** of the supplier's own products. Therefore, additional orders should only be accepted, if the own production flow will not be negatively influenced. Buyers cannot manufacture the products themselves or only with long waiting times, which is why the delivery time at the consumer side should be reduced to the initial situation.

The suppliers' **downtimes** should not be different compared to their own products. For the consumer, the situation can be different. The company can compensate the downtimes by placing orders externally.

Throughput times and **delivery reliability** behave in the same way as delivery times.
Machine utilization can be increased by participation in capacity platforms for the supplier. From the consumer perspective, capacity utilization is already very good, which is why no change is assumed.
The **personnel utilization** behaves simultaneously to the machine utilization.
The **inventories** itself will be higher from the supplier view. In the event of seasonal fluctuations, participation in capacity platforms can smooth out the inventories for the supplier, so that a constant level of inventories can exist throughout the year and the ordering cycle can remain constant. The consumer's inventories remain unchanged if no additional steps to the actual process is needed, except for the material, which is now used or available at the partner's company.
The supplier can optimize the **batch size** and thus indirectly optimize the production. An internal influencing factor is the company's own current order situation and the potential for combining internal and external orders into combined batches. For the consumer, the batch size does not change.
The last indicator is the **setup time**, which can increase for the supplier, but does not have to. The difference between in-house and external orders is not clear. Normally, no changes are expected for the consumer.

3.4 Application of the methodology and validation

A methodology was developed to support the decision-making process of companies, participating in a capacity sharing platform. This methodology can be used to support the utility of participation and the subsequent platform selection. In addition, changes in economic and logistic indicators were presented.
In the last step, this methodology will be transferred in a user-friendly application tool. In an Excel sheet, companies can specify their current situation according to the morphologies and receive information on the utility of participation, expressed as a percentage. An assignment to possible platforms will be provided in this Excel tool too. Further information on the changes in the indicators is also presented in this tool, depending on the initial scenario of a company. For this, the companies must choose a suitable scenario for their actual situation.
In addition to the implementation in an application, a guidance was developed. It presents the application and describes how to use it, as well as providing the basics of capacity sharing and further information on the platforms, which are available on the market. The companies can, without any research effort, recognize, whether participation is beneficial for them, as well as receive a pre-selection for possible platforms and view the expected changes of the economic and logistic indicators.

The methodology was developed with the input and discussions of experts, so that a constant scrutiny and validation of the sub steps has taken place. To validate the total methodology, the described questionnaire was sent to platform providers. The questionnaires were processed and returned from above 20 platform providers in a short time, which present the importance of this topic as well as a low-effort processing of the questionnaire. The manufacturing companies of the expert teams test the application afterwards. The results were discussed with the hole expert team, with the result, that in these cases, the methodology gives a good support. Further detailed validation steps must be done next.

4. Conclusion and outlook

The presented methodology as well as the implementation in an application and the provision of a guidance for the decision support, for or against a participation in capacity sharing, helps companies to deal with the topic and to be able to make individual decisions. The need for more flexibility due to order fluctuations is more important nowadays to survive in the market, and companies are aware of this. The market already provides several platforms that can be used. Nevertheless, there is currently a great deal of scepticism. Companies are unsure, whether participation is beneficial for them, which platform is suitable and what changes can be expected. This can be counteracted with the presented approaches. For the long-term use of these approaches, it is important to regularly check the platform market to include new market participants. In this way, the support can remain up to date and continue to help companies make decisions in the future.

* Funding notice

The IGF project 20801 N of the research association BVL was funded via the German Federation of Industrial Research Associations (AiF) in the programme of Industrial Collective Research (IGF) by the Federal Ministry for Economic Affairs and Climate Action (BMWK) based on a decision of the German Bundestag.

References

[1] Bundesministerium für Wirtschaft und Energie. Den digitalen Wandel gestalten. Online available https://www.bmwi.de/Redaktion/DE/Dossier/digitalisierung.html, 16.06.2021.

[2] Baums, A.; Schössler, M.; Scott, B. Kompendium Industrie 4.0. Online available https://docplayer.org/12198636-Kompendium-industrie-4-0-ansgar-baums-martin-schoessler-ben-scott-hg.html, 16.06.2021.

[3] Gracht, H.; Darkow, I.-L.; Hossenfelder, J. et al., 2010. Atmende Supply Chains. Wie gut ist Deutschlands gehobener Mittelstand auf volatile Märkte vorbereitet?.

[4] Müller, M.; Schüler, F.; Fritzsch, B. et al., 2020. Sharing von Produktionskapazitäten mittels digitaler Plattformen. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 115. Jg. (2020), H. 4, S. 257-261.

[5] Schwartz, M.; Gerstenberger, J., 2020. Corona-Krise im Mittelstand: Rückkehr zu voller Wirtschaftsaktivität in weiter Ferne, aber Lockerungen entspannen Liquidität. Online available https://www.kfw.de/PDF/Download-Center/Konzernthemen/Research/PDF-Dokumente-Fokus-Volkswirtschaft/Fokus-2020/Fokus-Nr.-294-Juni-2020-Corona-Mittelstand-2.pdf, 16.03.2022.

[6] Statista: Veränderung der Produktion im produzierenden Gewerbe zum Vormonat 2020-2021. Online available https://de.statista.com/statistik/daten/studie/150649/umfrage/produktion-im-produzierenden-gewerbe/, 16.06.2021.

[7] Heinen, T., 2020. Whitepaper Produktionsumfrage. Online available https://grean.de/wp-content/uploads/2020/07/Whitepaper_Fabriken-nach-Corona.pdf.

[8] KfW, 20212. Kurzfristiger Schock mit langfristiger Wirkung: Corona-Krise und internationale Wertschöpfungsketten.

[9] Spath, D.; Westkämper, E.; Bullinger, H-J. et al., 2017. Neue Entwicklungen in der Unternehmensorganisation, Springer vieweg, Berlin.

[10] Blickpunkt Arbeitsmarkt, 2022. Entwicklungen in der Zeitarbeit. Online available https://statistik.arbeitsagentur.de/DE/Statischer-Content/Statistiken/Themen-im-Fokus/Zeitarbeit/generische-Publikation/Arbeitsmarkt-Deutschland-Zeitarbeit-Aktuelle-Entwicklung.pdf?__blob=publicationFile, 16.03.2022.

[11] Leiss, F.; Wohlrabe, K., 2021. Aktuelle Entwicklungen bei Materialengpässen und Lieferproblemen in der deutschen Wirtschaft. Online available https://www.ifo.de/DocDL/sd-2021-digital-19-leiss-wohrabe-materialengpaesse.pdf, 16.03.2022.

[12] Voss-Dahm, D.; Mühge, G.; Schmierl, K. et al., 2011. Qualifizierte Facharbeit im Spannungsfeld von Flexibilität und Stabilität, Springer Fachmedien, Wiesbaden.

[13] Bornewasser, M.; Zülch, G., 2013. Arbeitszeit - Zeitarbeit. Springer Gabler, Wiesbaden.

[14] Eschberger, T. B2B Sharing: Der nächste Schritt für die Sharing Economy? Online available https://www.lead-innovation.com/blog/b2b-sharing, 16.06.2021.

[15] shareMag by MEWA. Vertrauen ist die Währung der Sharing Economy. Online available https://sharemag.de/vertrauen-ist-die-waehrung-der-sharing-economy, 16.06.2021.

[16] Schüler, F.; Petrik, D.,2021. Objectives of Platform Research: A Co-citation and Systematic Literature Review Analysis. In: Seiter, M.; Grünert, L.; Steur, A. (Hrsg.): Management digitaler Plattformen, Springer Gabler, Wiesbaden, S. 1-33.

[17] Kriwall, M.; Richter, J.; Fehlhaber, A. L. et al., 2018. Ganzheitliche Betrachtung der Ökologie und Logistikleistung von KMU. Wissenschaftliche Gesellschaft für Technische Logistik.

[18] Wegner, U.; Wegner, K., 2017. Einführung in das Logistik-Management. 3., aktualisierte und erweiterte Auflage, Springer Gabler, Wiesbaden.

[19] Gottmann, J., 2019. Produktionscontrolling. Springer Fachmedien, Wiesbaden.

[20] Herlyn, W., 2012. PPS im Automobilbau. 1. Aufl., Carl Hanser Fachbuchverlag.

[21] Nyhuis, P.; Wiendahl, H.-P. 2012. Logistische Kennlinien. 3. Auflage, Springer, Berlin, Heidelberg.

[22] Heiserich, O.-E.; Helbig, K.; Ullmann, W., 2011. Logistik. 4., vollst. überarb. und erw. Aufl., Gabler, Wiesbaden.

[23] proLogistik: Losgröße. Online available https://www.prologistik.com/logistik-lexikon/losgroesse/, 17.06.2021.

[24] Plinke, W.; Utzig, B. P., 2021. Industrielle Kostenrechnung. 9. Auflage, Springer Berlin; Springer Vieweg, Berlin.

[25] Namneck, A.; Böning, C.; Stonis, M., 2021. Reifegradbasierte Bewertung der Anforderungen einer erfolgreichen MES-Einführung. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, 116. Jg. (2021), H. 3, S. 175-179.

Biography

Maren Müller (*1993) is working at IPH – Institut für Integrierte Produktion Hannover gemeinnützige GmbH (non-profit limited company) as a project engineer in the field of logistics. In research and consulting projects she deals with storage planning and the selection and implementation of ERP- and ME-Systems.

Fabian Schüler (*1990) is working at IPRI – International Performance Research Institute gemeinnützige GmbH (non-profit limited company) as a Research assistant in the field of digital platforms. In research and consulting projects he deals with Industrial Internet of Things and the management of digital platforms.

Malte Stonis (*1979) is managing director of the IPH – Institut für Integrierte Produktion Hannover gemeinnützige GmbH (non-profit limited company). His research focuses on increasing the efficiency and digitization of production.

Peter Nyhuis (*1957) leads the Institute of Production Systems and Logistics (IFA) at the Leibniz Universität Hannover and is member of the management board of the IPH – Institut für Integrierte Produktion Hannover gemeinnützige GmbH (non-profit limited company).

3rd Conference on Production Systems and Logistics

Analyzing Supply Risks And Product Characteristics – A Systematic Literature Review

Maria Linnartz[1], Ursula Motz[1], Tobias Schröer[1], Volker Stich[1]
1Institut for Industrial Management (FIR) at RWTH Aachen, Aachen, Germany

Abstract

The environment in which companies operate is increasingly volatile and complex. This results in an increased exposure to disruptions. Past disruptions have especially affected procurement. Thus, companies need to prepare for disruptions. The preparedness for disruptions in the context of procurement is significantly influenced by the design of the procurement strategy. However, a high number of purchased articles and a variety of influencing factors lead to high complexity in procurement. The systematic design of the procurement strategy should therefore take into account the criticality of the purchased articles. This enables to focus on the purchased articles that have a high impact on the disruption preparedness. Existing approaches regarding the design of the procurement strategy in uncertain environments either lack practical applicability and objective evaluation or focus on the criticality of raw materials rather than of purchased articles. Therefore, a data-based approach for the systematic design of the procurement strategy in the context of the Internet of Production has been proposed. One central aspect of this approach is the identification of success-critical purchased articles. Thus, this paper proposes a framework for characterizing purchased articles regarding supply risks by combining two systematic analyses. First, a systematic literature review is performed to answer the question of what factors can be used to describe the supply risks of purchased articles. The results are analyzed regarding sources and impacts of risks and thus contribute to a structured characterization of supply risks. Second, existing criticality assessment approaches for raw materials are analyzed to identify categories and indicators that describe purchased articles. The results of both reviews provide the basis for linking product characteristics with supply risks and assessing product criticality which will be integrated into an app prototype.

Keywords

Disruptions; Supply Risks; Procurement Strategy; Product Characteristics; Internet of Production

1. Introduction

Past disruptions like the COVID-19 pandemic, the blockade of the Suez Canal or the flood in North-Rhine Westphalia have posed various challenges in supply chains and thus have demonstrated the need to prepare for disruptions. As past disruptions have demonstrated, the impact on the procurement side has been especially critical [1,2]. Procurement is responsible for organizing and ensuring the supply of external material and parts that are required for internal processes [1]. Preparing for disruptions should focus both on reducing the effect a disruption has on a company's performance as well as on enabling fast recovery after being disturbed. The preparedness for disruptions is significantly influenced by actions and strategic choices taken prior to a disruption [3]. For the area of procurement, the level of disruption preparedness is thus influenced by the design of the procurement strategy [1]. The procurement strategy determines the fundamental orientation and design of the supply process within the company and deals for example with

DOI: https://doi.org/10.15488/12195

the number of suppliers or the type of purchased objects [4]. A high level of complexity in procurement and a multitude of purchased products with different characteristics and various options for specifically shaping the procurement strategy contribute to the fact that one important prerequisite for increasing preparedness is having transparency of the current procurement situation and the purchased articles [5]. To handle the complexity, the systematic design of the procurement strategy with regards to disruptions should on one side take into account different data sources and on the other side consider the criticality of purchased articles. By taking into account the criticality of purchased articles, a focus can be set on articles with a high impact on preparedness. As part of the research project Cluster of Excellence "Internet of Production", existing approaches for developing procurement strategies and identifying critical purchased articles have been analyzed and it has been concluded that they either focus on raw materials or lack practical applicability and objective evaluation. Therefore, a framework for a data-based design of the procurement strategy in the context of the Internet of Production (IoP) has been proposed. [6]

This paper contributes to the framework and the criticality evaluation by systematically analyzing purchased articles regarding supply risks. On one hand, the paper aims at structuring supply risks by analyzing which aspects of supply risks are frequently mentioned in the literature and structuring them into distinct categories. On the other hand, the paper builds on existing work in the area of raw material criticality assessment to identify and structure indicators that characterize purchased articles. These sets of supply risk factors and purchased article characteristics are needed to link purchased articles to supply risks and assess their criticality. The systematic analysis thus provides the foundation for identifying relevant supply risks aspects based on different purchased article characteristics.

The remainder of this paper is organized as follows. Section 2 reviews the literature regarding supply risks and characterization of purchased articles. Section 3 presents the approach and the results of the systematic analyses. Section 4 summarizes the results and gives an outlook on the use of the results for building an app prototype.

2. State of the art

This section summarizes existing approaches to structure supply risks and to characterize purchased articles.

2.1 Supply risks

According to IVANOV AND SOKOLOV risks arise from uncertainty which is a general property of a system environment [7]. This understanding also underlies the definition by ROMEIKE. According to ROMEIKE, risks are possibilities to deviate from planned target values that result from the unpredictability of the future. [5] In the context of risks, this deviation is usually negative, while positive deviations are described as chances [8]. In the context of procurement, **supply risks** refer to "the uncertainty and severity of the events and consequences of any activity that adversely affects the inbound supply performance in terms of its target values" [2]. Supply risks affect the availability and quality of products that a company requires for its further value creation [9].

To analyze the current state regarding supply risks, general frameworks for structuring risks in supply chains are summarized first as they influence what risk factors are considered. In the literature, risks are categorized according to different criteria. SANCHIS AND POLER propose a general framework for disruption elements: A disruption is composed of a *source* that originates the disruption, a *disruptive event* that is the concrete incident that causes the negative effects and a *consequence* that is the impact of the disruptive event. [10]

Against this background, **cause-related categorizations** focus on different sources from which risks can occur. On a high level of abstraction, sources of risks in the context of supply chains can either be internal or external for the considered company or supply chain. [8] A common framework has been proposed by

MASON-JONES AND TOWILL which has been referred to by different authors [e. g. 11,8]. External risks are referred to as environmental risks. Internal risks can be further divided into company and supply chain internal risks. Supply chain internal risks include supply and demand risks while company internal risks are divided into process and control risks. SANCHIS AND POLER introduced a framework for origins of disruptions which summarizes and structures origins on different levels. On the first level, they distinguish between 11 origins: customers, distribution, economic/financial, energetic, environment, inventory, legislation, production, social, supply and technology. These are further divided into 59 sub-origins. [12]

Effect-related categorizations are oriented toward the consequences that result from the occurrence of risks. While operative risks refer to inherent fluctuations, disruptive risks are events that have a massive impact on the system. [8] SHEFFI ET AL. describe disruption in supply, disruption in transportation, disruption at facilities, freight breaches, disruption in communications and disruption in demand as possible failure modes within a supply chain [13]. CARVALHO identified four supply chain failures from the view of a single company: material shortage, capacity shortage, finished product completed but not delivered and scrap/rework [14]. With a focus on supply risks, VON CUBE ET AL. propose deviations from the expected quality of delivered parts, deviations from the scheduled date of delivery of ordered lots, deviations from the planned quantity of delivered goods, and deviations from the planned prices of procured goods as possible consequences of supply disruptions [15]. SUCKY distinguishes supply risks in quality, quantity, timing, cost and transport risks where deviations can occur [9].

Additionally, there are literature reviews that summarize different risks factors that affect several of the above-mentioned categories. One example is the review presented by HO ET AL. who identified various risks factors in the context of supply chains. These factors are classified into macro risks, demand risks, manufacturing risks, supply risks and infrastructure risks (information, transportation and financial risks). [16] HUNDNURKAR ET AL. propose a supply chain risk classification scheme that distinguishes between risk sources and risks. The classification uses product characteristics, supply chain management processes, supply chain infrastructure, external environment and human resources as categories. [17] HOFFMANN AND ROLAND differentiate supply risks into environmental and behavioral risks. The behavioral risks are further divided into financial, operative, and strategic risks. For each category, they summarize relevant risks factors. [18] WIEDENMANN AND GRÖßLER focus on the identification and categorization of relevant supply risks in manufacturing supply networks. They use a mixed-method approach that combines a structured literature review and interviews. The proposed framework contains six risk dimensions (quality, delivery, collaboration, economic, ambience and compliance) which refer to the outcomes of risks and 27 supply risks factors that refer to the sources of risks. [2]

This analysis demonstrates that the literature in terms of supply chain and supply risks is complex due to the use of different terminologies (e. g. risks, disruptions, disturbances, vulnerability) and various levels of consideration. Moreover, a vast amount of categorization approaches exists that each focus on different aspects when defining main and subcategories. The approaches often do not differentiate between sources of risks, disturbance events and consequences which increases the complexity and complicates structured analysis. Additionally, only a few approaches explicitly focus on supply risk categorization and specification. Thus, a structured analysis of supply risks is needed as a basis for a systematic design of procurement strategies.

2.2 Characterization of purchased articles

As described in a previous publication in the context of this research project, portfolio methods are often used to cluster purchased articles and derive procurement strategies. These approaches contain different indicators to characterize purchased articles regarding their supply risks. They typically divide the indicators considered into an external and an internal dimension. A limitation of portfolio methods is the lack of objectivity in the evaluation of the indicators used. A data-based evaluation is often missing. [6]

Additionally, these approaches mostly do not categorize the indicators. Due to a lack of criticality assessments for purchased articles, raw material criticality assessments were analyzed. It was found, that existing assessment approaches focus on different risks and indicators as well as on different aggregation levels [19]. Selecting the right indicators thus poses a challenge. [6] As stated before, the extent to which the identified indicators are applicable for the assessment of purchased articles other than raw material needs to be examined. These approaches thus serve as an input for the structuring of the purchased article characteristics in the next section.

3. Framework for analyzing supply risks

The proposed framework is part of the approach for the data-based design of procurement strategies in the IoP that was introduced by the authors in previous work [6]. In this approach, action research and the CRISP-DM model are combined. By going through different action research cycles, the approach aims at ensuring the practical applicability and fast implementation of the results in companies. This paper contributes to the first action research cycle, which aims at characterizing purchased articles in the context of supply risks. It lies the foundation for the second cycle which concentrates on implementing a calculation logic that enables the identification of success-critical purchased articles. This paper focuses on characterizing supply risks and purchased articles. The systematic analysis of both areas results in two factor catalogues and is the foundation for the structured analysis of interdependencies between the developed catalogues.

3.1 Structuring of supply risks factors

In this section, the results for structuring supply risks factors are presented. The results are based on a systematic literature review. This research method is chosen since systematic literature reviews offer the possibility to integrate different findings and perspectives and create an overview of already existing empirical evidence. By using a systematic approach for the literature analysis reliable findings and conclusions can be generated while bias can be minimized. [20] Systematic literature reviews are thus replicable and transparent [21]. This research uses the five-step approach for conducting systematic literature reviews in management and organization studies which has been proposed by DENYER AND TRANFIELD. The systematic analysis starts by formulating a research question (*Step 1*) which is followed by locating the studies (*Step 2*). Locating the studies includes the selection of databases and search algorithms. This step is followed by studying and evaluating the studies (*Step 3*). To do so, criteria for inclusion and exclusion of studies need to be defined. These selection criteria must be recorded to meet the requirement of transparency. After studies have been selected, the analysis and synthesis take place (*Step 4*). While the analysis focuses on identifying the individual and constituent parts, the synthesis aims at finding associations between the identified parts. Synthesis thus goes beyond pure description and aims to create knowledge through combining different studies. The final step is the reporting and use of the results (*Step 5*). [22]

Based on the underlying approach and the state of the art regarding supply risks, this literature review aims at answering the question "*What factors can be used to describe the supply risk of purchased products?*". The results of the research contribute to the structuring of different supply risk dimensions and corresponding negative effects on manufacturing companies. In this work, supply risk factors are understood as potential sources of supply disruptions that can cause different kinds of impacts.

To identify relevant studies and cover a range of different types of information, various keywords were identified and subsequently a range of search strings was developed. Several keywords for the context of procurement *(supply*, *upstream*, *supplier*, *procurement*, *source*) were combined with keywords related to risks (*risk*, *disruption*, *disturbance*, *critical*, *vulnerable*) to identify risks that occur within procurement. Additionally, keywords were added to address any existing frameworks and measurement approaches (*framework*, *assessment*, *evaluation*, *measurement*, *identification*, *classification*) as well as keywords that focus on factors (*characteristic*, *feature*, *indicator*, *parameter*, *category*). Scopus was chosen as a search database as it covers a wide range of articles. The search was performed in September 2021 and was limited to articles written in English. Additional search conventions such as filters for subject areas (*Business, Management and Accounting*; *Decision Sciences*; *Engineering*) have been added. After this step, a total of 554 studies were found and were further examined in the third step of the process.

Titles, keywords and abstracts of the articles were read to determine their suitability for inclusion. Studies were included if they focused on supply risks and identified specific factors that detail these risks. Studies that did not use the supply of a production firm as their primary focus, such as articles describing energy procurement, agricultural procurement or healthcare procurement were excluded as the identified factors were too specific. This process retained 142 studies out of 554 studies. Afterwards, the full texts of the studies were closely examined, removing another 100 studies as they only marginally addressed risk factors in the procurement context and therefore did not have an additional value for the paper. Particular attention has been given to studies that identify risk factors in the supply context and attempt to provide a framework, resulting in a total of 42 studies to be considered for analysis and synthesis. Figure 1 summarizes the successive reduction of the relevant studies.

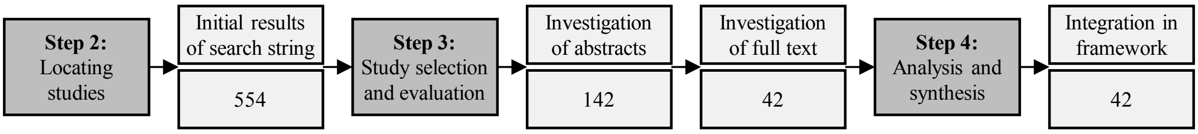


Figure 1: Progress of the systematic literature review and successive reduction of relevant studies

During the fourth step, each study was first analyzed concerning the supply risk factors it contained. The supply risks factors identified in these papers were extracted. It was striking during the following analysis that many studies chose different levels of classification and thus no clear risk assignment has been established so far. Furthermore, the wording of supply risk factors was inconsistent, resulting in a blending of risk factors, disruption events, supplier or product characteristics, and disruption consequences. The result of the analysis of the selected paper was an unstructured list of factors related to supply risks as a basis for the following compilation. Next, the wording of the factors was adapted based on their description to reflect their underlying meaning. This enabled the aggregation of factors and the removal of factors that did not specifically refer to the procurement context. Moreover, factors that did not represent risks, but characteristics or consequences were identified. These were not included in the proposed framework. Considering, existing classification schemes, the identified risk factors were grouped into five categories. The categories and risk factors es are summarized in Figure 2.

Supplier risks	**Collaboration risks**	**Transportation risks**	**Supply market risks**	**Ambience risks**
• Capacity constraints • Volume flexibility • Labor practice • Location • Technical capability • Quality requirements capability • Delivery requirements capability	• Information availability • Opportunistic behavior • Lock-in effect	• Transport capacity constraints • Handling capability • Transport failure • Transport complexity • Transport restrictions	• Input material availability • Supply sources availability • Supply market development	• Geopolitical conditions • Natural hazards • Man-made incidents • Health issues • Economic issues

Figure 2: Framework for supply risk factors

Supplier risks contain factors that arise from suppliers and relate to a specific supplier. Risk factors with regards to the production capacity of a supplier are capacity constraints [e. g. 23,24] and volume flexibility [e. g. 24,25]. They can lead to problems in the context of material availability. Besides, labor practice [e. g. 26,27] can pose supply risks, as unethical practices like child labor can result in production restrictions. The supplier location [e. g. 24,28] contributes to site-specific risks, which are closely linked to the environmental risks discussed later on. Additionally, the capability of a supplier with regards to technical [e. g. 25,28], quality [e. g. 29,26] and delivery issues [e. g. 29,26,30] are important risks factors. Technical capability refers to the technical skills that are needed to produce the articles, while quality requirements capability relates to the processes and control to ensure the quality of products. Delivery requirements capability is the ability to deliver the requested article at the requested time.

Collaboration risks refer to the relationship with a supplier. An important risk factor mentioned by various authors is the information availability [e. g. 28,29]. Problems occur when information is delayed or the communication networks are unstable. The behavior within a collaboration poses a risk if one partner takes opportunistic actions [e. g. 26,30]. Another risk factor in this context is the lock-in effect [e. g. 31,27]. In that case, switching suppliers is difficult which hinders flexibility.

Transportation risks summarize factors that relate to the logistical aspects of procurement. Risk factors like transport capacity constraints [e. g. 32,31] and handling capability [e. g. 32,33] correspond to the factors of the category supplier risks. Capacity constraints include shortages in terms of a specific transport mode or shortages in space and containers. The factor handling capability refers to the required ability when packing, loading and transporting goods so that no damages occur. Additionally, transport failure [e. g. 29,34], transport complexity [e. g. 24,35] and transport restrictions [e. g. 24,30] are relevant risk factors in this context. Transport failure refers to the impossibility of transport execution for example due to an accident or transportation breakdown. Transport complexity is influenced for example through the number of transfer points and can increase vulnerability. Transport restrictions can occur when crossing borders are required and can impact timely delivery.

Supply market risks focus on the supply market and its development as a whole. Risk factors relate to the availability of input material [e. g. 25,2] and supply sources [e. g. 36,26]. The factor input material availability refers to the availability of raw material and the input that is required by the suppliers. The factor supply sources availability includes the existence of potential suppliers. Another risk factor is the supply market development [e. g. 36,37], which includes for example volatility in terms of the number of consumers and suppliers.

Ambient risks contain risk factors that arise from the supply environment. Following WIEDENMANN AND GRÖßLER this category is named ambient risks to avoid ambiguous interpretation [2]. This category is broader than the mere supply market risks and contains factors that cannot be influenced as such and affect various areas in the supply environment. Risk factors are geopolitical conditions [e. g. 35,34], natural hazards [e. g. 35,34], man-made incidents [e. g. 30,2] as well as health issues [e. g. 27,37] and economic issues [e. g. 35,33]. Geopolitical conditions are for example influenced by political conflicts and unrest or war but also include export or import restrictions that hinder supply activities. Natural hazards include tsunamis or earthquakes while terrorism is an example of man-made incidents. The factor health issues contains for example pandemics. Economic issues include currency fluctuations, stock market instability, global economic performance and inflation. [35]

Following VON CUBE ET AL., impacts of supply risks can be distinguished in deviations from the expected quality, the planned quantity, the scheduled date of delivery and the planned price [15]. As an additional impact, the overall availability of the procured material was added to the framework as it results in deviations regarding quantity, time and price. Each risk factor can be linked to at least one of these impacts.

3.2 Structuring of purchased article characteristics

This section presents the results regarding the structuring of purchased article characteristics. In previous research by the authors existing criticality assessment approaches for raw materials have been identified [6]. These approaches as well as studies from the systematic literature review described above are the basis for the following analysis. The aim is to identify indicators and categories that describe purchased articles in the context of supply risks.

Previous research analyzed eight different approaches, dealing with criticality assessment approaches for raw materials [38–44,19]. For this paper, each study was re-examined, and the indicators used to describe and categorize the purchased articles were incorporated into an unstructured list. Next, it was examined which studies explicitly referred to raw materials and whether they applied to the broader category purchased articles as well. Most characteristics were applicable if they did not specifically relate to raw material extraction in terms of mining. Besides, two other sources that characterize purchased articles were considered in the analysis [45,46]. In addition to these sources, results from the systematic literature review were taken into account. During the systematic literature review, it was found that some of the identified risk factors were not related to the general context of the procurement process, but were very specific in terms of the purchased articles. With regard to the structuring of purchased article characteristics, the risks and indicators directly relating to product characteristics were not considered in the above-described framework but analyzed here. Following the above-described approach for the synthesis and analysis, the characteristics were collected in a list and their wording was adjusted to reflect the underlying meaning. This enabled the aggregation of the named characteristics. This was followed by a grouping and categorization of the characteristics of the purchased articles. The compiled indicators needed to characterize the product and at the same time can be linked to the previously listed supply risks. As can be seen in Figure 3, four categories were established. The characteristics are either not-related to suppliers (product characteristics and economical aspects) or related to suppliers (supplier characteristics and logistical aspects).

Product characteristics	Economical aspects	Supplier characteristics	Logistical aspects
• Product specialization • Product vulnerability • Hazard risk / safety specifications • Frequency of product changes • Product's lifecycle position • Demand volatility • Resource competition • Substitution possibility	• Volume purchased • Total purchasing cost • Material utilization • Material value • Strategic importance • Revenue impact • Price volatility	• Number of suppliers • Diversity of supply • Lead time • Reliability	• Import dependency • Supply distance

Figure 3: Framework for purchased articles characteristics

Product characteristics cover not only the physical characteristics of the purchased product but also relate to its specific supply market. Product specialization [e. g. 44,46] describes the uniqueness of a purchased article in terms of its level of complexity and individuality. Product vulnerability [e. g. 46] describes the susceptibility of a procured article to external influences that have a diminishing effect on its performance level. The hazard risk or safety specification of a product [e. g. 23,26] is based on its occupational requirements in terms of physical, chemical, biological or ergonomic requirements. The frequency of product changes [e. g. 23,36,46] describes the number and intensity of both technical and design changes to the purchased article. A product's lifecycle position [e. g. 45,26] is usually defined by how long it has already been on the market and whether there is a chance that it might be discontinued. One characteristic that is also frequently mentioned regarding supply risks is the volatility of the demand [e. g. 24,43,19]. Resource competition [e. g. 42,43] refers to the popularity of the article and is based on competing demand for an article that is available only in limited quantities. Substitution possibility [e. g. 42,46] describes how well the purchased article can be replaced through another article.

Economical aspects are mainly intended to describe the importance of an article for the buying company and the effects on economic targets. The purchasing volume or consumption volume [e. g. 40–42] is the quantity of a purchased article that is ordered within a certain period. Purchasing costs [e. g. 45,41] are understood to be the total costs for carrying out the procurement process, so that the importance of an article is described by its percentage of the total purchasing costs [42]. Material utilization [e. g. 45,47] describes the extent to which a purchased article is used in finished products. The material value [45,40] refers to the financial value of the utilized material and is directly related to the purchasing cost. Strategic importance characterizes the impact of a product on the company's strategic objectives [e. g. 40,42], while revenue impact [e. g. 19,46] expresses the importance of the product to the company's income. Price volatility [e. g. 39,41] describes the tendency of a product to fluctuate in price, which is explained, for example, by changes in the market or regular changes in the price of related raw materials. The tendency of a product to develop price spikes is also considered here.

A purchased article is moreover characterized through the suppliers that deliver this article. **Supplier characteristics** thus contain characteristics that relate to the actual suppliers from which a company purchases the articles. The number of suppliers [e. g. 26,42] is the actual amount of available and capable suppliers. Despite the number of suppliers, a product is characterized through the diversity of supply [e. g. 48,19] which refers to the geological distribution of suppliers and their production or export structures. A purchased article is also characterized through the lead time needed for its delivery [e. g. 45,39,28] which is dependent on the supplier. The last characteristic that directly refers to a supplier is its reliability [e. g. 45,39]. It includes the adherence to delivery dates, quality and delivery amount aspects.

As the last category, **logistical aspects** include characteristics that refer to the transportation process which is performed when delivering the articles. These characteristics are thus influenced by the location of the suppliers. Import dependency [e. g. 40,19] indicates the extent to which an article must be imported or is available domestically. Supply distance [e. g. 41,46] refers to the transport distance between the suppliers and the location of the buying company. The supply distance influences possible modes of transportation, delivery time and transportation costs.

4. Conclusion and outlook

Risks in the context of procurement are interpreted and understood in various ways within the literature. The systematic design of the procurement strategy influences the risk exposure and thus requires knowledge about the relevant risk factors. Additionally, complexity in procurement requires focusing on articles with a high impact on disruption preparedness. Therefore, supply risks factors and purchased article characteristics have been systematically studied. Using a systematic literature analysis, a variety of sources were analyzed. Based on the results a framework for supply risk factors was proposed which contains five categories and 23 risk factors. The identified risk factors are potential sources of supply disruptions. The disruptions can result in deviations from the expected quality, the planned quantity, the scheduled date of delivery, the planned price and deviations in the overall availability of the procured material. Additionally, purchased article characteristics were analyzed. The resulting framework for purchased article characteristics includes characteristics related to and not related to suppliers and is structured in four categories with 21 characteristics. The results create transparency on the relevant aspects both in the context of supply risks and purchased article characteristics. They are thus the basis to analyze the interdependencies between article characteristics and supply risks. The supply risks for an article then result from a combination of certain purchased article characteristics. This allows the identification of critical articles. Further research is needed to identify the links between the identified factors and characteristics. In this context, interdependencies between different risk factors should be taken into account as well. Once the links have been analyzed the results will be integrated into an app prototype which enables the identification of critical purchased articles. The app prototype will use data from different business application systems to characterize purchased articles and derive statements regarding their supply risks.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Pereira, C.R., Christopher, M., Lago da Silva, A., 2014. Achieving supply chain resilience: the role of procurement. Supply Chain Management: An International Journal 19 (5/6), 626–642.

[2] Wiedenmann, M., Größler, A., 2020. Supply risk identification in manufacturing supply networks. The International Journal of Logistics Management.

[3] Kiebler, L., Ebel, D., Klink, P., Sardesai, S., 2020. Risikomanagement disruptiver Ereignisse in Supply Chains. Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund, 20 pp. https://www.iml.fraunhofer.de/content/dam/iml/de/documents/OE%20220/Risikomanagement_disruptiver_Ereignisse_in_Supply_Chains.pdf. Accessed 18 March 2021.

[4] Lasch, R., 2019. Strategisches und operatives Logistikmanagement: Beschaffung, 2., überarbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden Germany, 326 pp.

[5] Romeike, F., Hager, P., 2020. Risiko-Management in der Logistik und Supply Chain, in: Romeike, F., Hager, P. (Eds.), Erfolgsfaktor Risiko-Management 4.0. Methoden, Beispiele, Checklisten Praxishandbuch für Industrie und Handel, 4., vollst. überarb. Auflage 2020 ed. Springer Fachmedien Wiesbaden GmbH; Springer Gabler, Wiesbaden, pp. 353–385.

[6] Linnartz, M., Motz, U., Schröer, T., Stich, V., Müller, K., Greb, C., 2021. Increasing Resilience in Procurement in the Context of the Internet of Production. Journal of Production Systems and Logistics 1 (2021).

[7] Ivanov, D., 2018. Structural Dynamics and Resilience in Supply Chain Risk Management. Springer International Publishing, Cham.

[8] Zitzmann, I., 2018. Supply Chain-Flexibilität zur Bewältigung von Unsicherheiten. Dissertation. University of Bamberg Press, Magdeburg.

[9] Sucky, E., Zitzmann, I., 2016. Risikomanagement in nachhaltigen Supply Chains, in: Eckert, S., Trautnitz, G. (Eds.), Internationales Management und die Grundlagen des globalisierten Kapitalismus. Springer Fachmedien Wiesbaden, Wiesbaden, 459-478.

[10] Sanchis, R., Poler, R., 2014. Enterprise resilience assessment: a categorisation framework of disruptions. Dirección y Organización 54, 45–53.

[11] Christopher, M., Peck, H., 2004. Building the Resilient Supply Chain. The International Journal of Logistics Management 15 (2), 1–14.

[12] Sanchis, R., Poler, R., 2019. Origins of Disruptions Sources Framework to Support the Enterprise Resilience Analysis. IFAC-PapersOnLine 52 (13), 2062–2067.

[13] Sheffi, Y., Caniato, F., Rice, J., 2003. A Supply Chain Response to Global Terrorism: A Situation Scan, 10 pp. Accessed 22 November 2021.

[14] Carvalho Remigio, H., 2012. Modelling resilience in supply chain. Dissertation, Lissabon, 196 pp.

[15] Cube, J.P. von, Härtel, L., Schmitt, R., 2016. Model-based Decision Support in Supply Chains – Requirements for Monetary Supply Risk Quantification. Procedia CIRP 57 (7), 171–176.

[16] Ho, W., Zheng, T., Yildiz, H., Talluri, S., 2015. Supply chain risk management: a literature review. International Journal of Production Research 53 (16), 5031–5069.

[17] Hudnurkar, M., Deshpande, S., Rathod, U., Jakhar, S.K., 2017. Supply Chain Risk Classification Schemes: A Literature Review. Operations and Supply Chain Management: An International Journal (13), 182–199.

[18] Hoffmann, S., Roland, F., 2013. Nutzungspotenziale von Social Media im Supplier Risk Management, in: Bogaschewsky, R. (Ed.), Supply Management Research. Aktuelle Forschungsergebnisse 2013, 1st ed. ed. Springer Fachmedien Wiesbaden GmbH, Wiesbaden, pp. 223–250.

[19] Schrijvers, D., Hool, A., Blengini, G.A., Chen, W.-Q., Dewulf, J., Eggert, R., van Ellen, L., Gauss, R., Goddin, J., Habib, K., Hagelüken, C., Hirohata, A., Hofmann-Amtenbrink, M., Kosmol, J., Le Gleuher, M., Grohol, M., Ku, A., Lee, M.-H., Liu, G., Nansai, K., Nuss, P., Peck, D., Reller, A., Sonnemann, G., Tercero, L., Thorenz, A., Wäger, P.A., 2020. A review of methods and data to determine raw material criticality. Resources, Conservation and Recycling 155, 104617.

[20] Snyder, H., 2019. Literature review as a research methodology: An overview and guidelines. Journal of Business Research 104 (5), 333–339.

[21] Tranfield, D., Denyer, D., Smart, P., 2003. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. British Journal of Management 14 (3), 207–222.

[22] Denyer, D., Tranfield, D., 2011. Producing a systematic review, in: Buchanan, D.A., Bryman, A. (Eds.), The Sage handbook of organizational research methods. Sage Publications Inc, Thousand Oaks, CA, pp. 671–689.

[23] Aqlan, F., Lam, S.S., 2015. A fuzzy-based integrated framework for supply chain risk assessment. International Journal of Production Economics 161, 54–63.

[24] Blackhurst, J.V., Scheibe, K.P., Johnson, D.J., 2008. Supplier risk assessment and monitoring for the automotive industry. International Journal of Physical Distribution & Logistics Management 38 (2), 143–165.

[25] Govindan, K., Jepsen, M.B., 2016. Supplier risk assessment based on trapezoidal intuitionistic fuzzy numbers and ELECTRE TRI-C: a case illustration involving service suppliers. Journal of the Operational Research Society 67 (2), 339–376.

[26] K.T., R., Sarmah, S.P., Tarei, P.K., 2019. An integrated framework for the assessment of inbound supply risk and prioritization of the risk drivers. Benchmarking: An International Journal 27 (3), 1261–1286.

[27] Shahbaz, M.S., RM Rasi, R.Z., Bin Ahmad, M.F., 2019. A novel classification of supply chain risks: Scale development and validation. Journal of Industrial Engineering and Management 12 (1), 201.

[28] Guertler, B., Spinler, S., 2015. Supply risk interrelationships and the derivation of key supply risk indicators. Technological Forecasting and Social Change 92, 224–236.

[29] Junaid, M., Xue, Y., Syed, M.W., 2020. Construction of index system for risk assessment in supply chains of automotive industry. International Journal of Supply Chain Management 9 (4), 91–106.

[30] Kumar Sharma, S., Sharma, S., 2015. Developing a Bayesian Network Model for Supply Chain Risk Assessment. Supply Chain Forum: An International Journal 16 (4), 50–72.

[31] Lockamy III, A., 2014. Assessing disaster risks in supply chains. Industrial Management & Data Systems 114 (5), 755–777.

[32] Kwak, D.-W., Rodrigues, V.S., Mason, R., Pettit, S., Beresford, A., 2018. Risk interaction identification in international supply chain logistics. International Journal of Operations & Production Management 38 (2), 372–389.

[33] Manikandan, L., Thamaraiselvan, N., Punniyamoorthy, M., 2011. An instrument to assess supply chain risk: establishing content validity. International Journal of Enterprise Network Management 4 (4), 325.

[34] Paksoy, T., Çalik, A., Yildizbaşi, A., Huber, S., 2019. Risk Management in Lean & Green Supply Chain: A Novel Fuzzy Linguistic Risk Assessment Approach, in: Paksoy, T., Weber, G.-W., Huber, S. (Eds.), Lean and Green Supply Chain Management, vol. 273. Springer International Publishing, Cham, pp. 75–100.

[35] Burns, L., 2017. An extended framework for supply chain risk management: incorporating the complexities of emerging industries and large-scale systems. International Journal of Manufacturing Technology and Management 31 (1/2/3), 217.

[36] Ellis, S.C., Shockley, J., Henry, R.M., 2011. Making Sense of Supply Disruption Risk Research: A Conceptual Framework Grounded in Enactment Theory. Journal of Supply Chain Management 47 (2), 65–96.

[37] Shi, Y., Zhang, Z., Wang, K., 2017. A Dempster Shafer Theory and Fuzzy-Based Integrated Framework for Supply Chain Risk Assessment, in: Uden, L., Lu, W., Ting, I.-H. (Eds.), Knowledge Management in Organizations, vol. 731. Springer International Publishing, Cham, pp. 347–361.

[38] Dixit, V., 2020. Risk assessment of different sourcing contract scenarios in project procurement. International Journal of Construction Management, 1–13.

[39] Griffin, G., Gaustad, G., Badami, K., 2019. A framework for firm-level critical material supply management and mitigation. Resources Policy 60, 262–276.

[40] Helbig, C., Wietschel, L., Thorenz, A., Tuma, A., 2016. How to evaluate raw material vulnerability - An overview. Resources Policy 48, 13–24.

[41] Jaipuria, S., Jenamani, M., Ramkumar, M., 2016. The strategic procurement of raw material: a case study. International Journal of Procurement Management 9 (5), 524–547.

[42] Kraljic, P., 1983. Purchasing Must Become Supply Management. Harvard Business Review 61 (5), 109–117.

[43] Lapko, Y., Trucco, P., Nuur, C., 2016. The business perspective on materials criticality: Evidence from manufacturers. Resources Policy 50, 93–107.

[44] Montgomery, R.T., Ogden, J.A., Boehmke, B.C., 2018. A quantified Kraljic Portfolio Matrix: Using decision analysis for strategic purchasing. Journal of Purchasing and Supply Management 24 (3), 192–203.

[45] Bauhoff, F., 2013. Selbstoptimierende Regelung der artikelbezogenen Materialdisposition in der Beschaffung. Zugl.: Aachen, Techn. Hochsch., Diss., 2013, 1. Aufl. ed. Apprimus Verlag, Aachen, 235 pp.

[46] Spille, J., 2009. Typspezifisches Risikomanagement für die Beschaffung von Produktionsmaterialien in der Automobilzulieferindustrie. Zugl.: Aachen, Techn. Hochsch., Diss., 2008. Shaker, Aachen, 252 pp.

[47] Ellis, S.C., Henry, R.M., Shockley, J., 2010. Buyer perceptions of supply disruption risk: A behavioral view and empirical assessment. Journal of Operations Management 28 (1), 34–46.

[48] Helbig, C., Bruckler, M., Thorenz, A., Tuma, A., 2021. An Overview of Indicator Choice and Normalization in Raw Material Supply Risk Assessments. Resources 10 (8), 79.

Biography

Maria Linnartz, M. Sc. (*1993) is a research associate at the Institute for Industrial Management (FIR) at the RWTH Aachen in the research group Supply Chain Management. She studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. Currently, she is working on the research project "Internet of Production", which aims to develop data-driven decision support systems.

Ursula Motz, B. Sc. (*1996) studies Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. She is a research assistant at the Institute for Industrial Management (FIR) at the RWTH Aachen.

Tobias Schröer, M. Sc. (*1991) studied Business and Engineering at the TU Clausthal. Since 2016, he has been working as a research assistant at the Institute for Industrial Management (FIR) at RWTH Aachen and is head of the Production Management department since 2020.

Prof. Dr.-Ing. Volker Stich (*1954) has been CEO of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and led the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

General Approach And Prerequisites For Transferring Factory Planning Methods On Flow Orientation And Transformability To Hospital Systems

Lena Wecken[1], Shiva Faeghi[2], Lennart Hingst[1], Peter Nyhuis[1], Kunibert Lennerts[2]
[1]*Institute of Production Systems and Logistics, Garbsen, Germany*
[2]*Institute of Technology and Management in Construction, Karlsruhe Instutute of Technology, Karlsruhe, Germany*

Abstract

Manufacturing companies are faced with the challenge of operating cost-efficiently and remaining competitive in a turbulent environment with constantly changing demands on production. To meet these challenges, factory planning has developed concepts of flow orientation and transformability. Through decades of research, factory planners now have extensive methodologies and numerous principles, enabling them to design factory objects appropriately and align factories to be flow-oriented and transformable. Hospitals face similar challenges like manufacturing companies. Due to public funding, many hospitals have limited financial resources and must, for example, cover parts of their financial requirements by themselves through cost-efficiency. In addition, hospitals are influenced by ongoing developments like demographic change and recent challenges such as the COVID-19 pandemic. These and further examples not only tighten the economic situation of hospitals, but also force their systems to adapt to resulting challenges. They must successfully align their systems to remain operational under changing conditions. In contrast to factories, these issues have not been addressed sufficiently in the field of hospital planning. Therefore, factory planning approaches on flow orientation and transformability will be transferred to hospital systems in order to strengthen hospitals against globally existing and socially relevant challenges in the healthcare system. With the aim to realise this venture, this paper presents a structured approach for its implementation. It also investigates the fundamental similarities between factories and hospitals and examines whether the main prerequisites for the successful transfer of the approaches can be met.

Keywords

Factory Planning; Flow Orientation; Transformability; Hospital Planning; Hospital System

1. Introduction and need for research

Requirements for hospital systems have increased over time. This applies to those resulting from both external influences and internal properties. In Germany, efforts to economise hospitals were made with the amendment of the "Law on the Reform of Statutory Health Insurance" in 2000 [1] and was accompanied by increasing cost pressure **(external influence)** on hospitals. Hospitals are now only paid a flat rate per case according to Diagnosis Related Groups (DRG), instead of being reimbursed for the individually incurred costs. This forces hospitals to strive for economic efficiency in order to cover the costs of treatment in every case [2]. To be able to cover future investment needs independently, hospitals must also generate a profit by keeping their costs significantly below the case-based flat rates. This is necessary since federal states provide increasingly less funding in the health system [3,4]. However, operational business can only be economical if the processes run efficiently. Hospitals are characterised by a function-orientation where departments have

DOI: https://doi.org/10.15488/12126

a pronounced divisional thinking. Hence, the hospital system can be seen as a federal system comprised of different and independent but incompatible organisations. This leads to interface problems between different disciplines or departments and is also noticeable in hospital layouts by long travel and waiting times. Thus, a cross-departmental, more efficient form of service provision is not realisable [4-7]. Due to the process inefficiencies **(internal property)** many hospitals are unable to meet the demand for economic operations.

In addition to the above-mentioned requirements, hospitals must also cope with change drivers **(external influence)** resulting from their turbulent environment [6]. Change drivers can have many different origins. Examples include demographic changes and the shift in the distribution of disease patterns [8], the consolidation of the hospital market [9] or, most recently, the COVID-19 pandemic [10]. Hospital systems must be able to adapt to these change drivers in order to remain operational in the short term and competitive in the medium to long term. Examples from practice show that hospitals cannot meet this requirement due to their transformation inertia **(internal property)** [3,4,6,11,12].The change drivers and the changes resulting from them must be identified at an early stage, which applies not only to hospital operations but already to hospital construction planning [13], which are characterised by long planning periods. There is a high risk that they will no longer meet the requirements by the time they are put into operation [12]. Therefore, it is necessary, that the consequences can be mitigated at an early stage with appropriate measures, such as accelerated reutilisation cycles in construction [14].

As summarised in Figure 1, the performance of hospital systems is strongly affected by the external influences of cost pressure and change drivers. The system properties of process inefficiencies and transformation inertia of hospitals do not provide an opportunity to adequately address the external influences, but rather worsen the situation with an additional negative impact on performance.

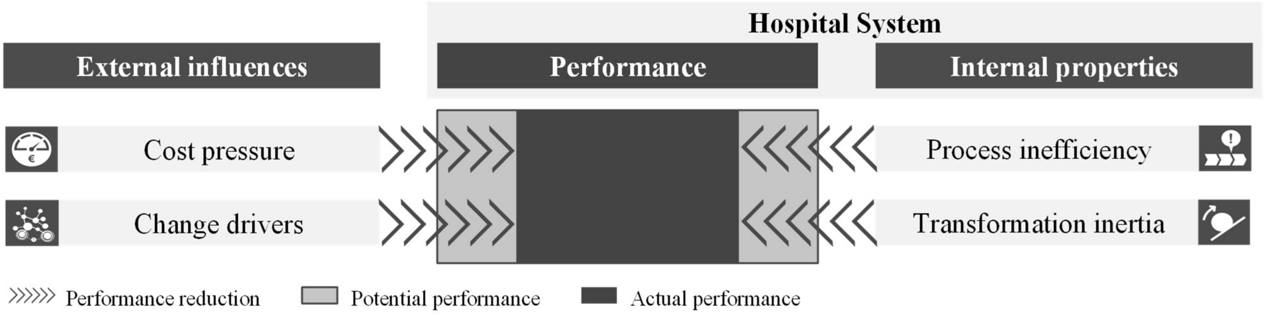


Figure 1: External Influences and Internal Properties of Hospital Systems

To increase the process efficiency, the ideal layout of a hospital must be aligned with the treatment processes to reduce travel and waiting times. In addition, the hospital system must be designed in such a way that it is prepared against the change drivers. Consequently, it requires the integration of flow orientation, i.e. the orientation towards process-oriented structures [15], and transformability, i.e. the ability to realise changes beyond defined areas and are absolutely necessary to ensure future oriented developments [16]. These two features enable the hospitals to shape the internal properties and cope adequately with external influences. Hospital planning, here also refers to hospital construction planning, has not been properly scrutinised as an object of research for decades. Hence, the methodologies in this field are lagging behind [13,17]. Therefore, there has been no approach that enables flow orientation and transformability equally important with the necessary depth of detail for hospital planning.

Regarding process efficiency, the approaches aim to reduce the effort required to provide services. For example, they maximise the utilisation of investment and operating cost-intensive resources by centralising functional areas [12] or minimise process times through lean methods [19]. The layout planning has a great influence on the process times through the determination of the travel times, whereby a process analysis for flow-oriented layout planning does not take place in the approaches examined in the current study. Existing approaches only consider the processes in the context of process management [2], during the introduction of activity-based costing [20,21] or for the implementation of IT systems [7,22]. Moreover, numerous quantitative approaches have been developed for supporting decision-making in layout planning, which arrange the spaces in individual hospital areas based on objective functions under consideration of certain constraints. The approaches considered can be divided into mathematical optimisation methods [5,23-26], simulation models [27–32] and hybrid approaches [33–36].

Regarding flexibility, existing approaches understand this term as a requirement that is becoming increasingly important due to the permanent change in the healthcare sector [6]. Flexible hospitals are recognised worldwide and are characterised, among other things, by decentralised structures [12,37], additional capacities [38,39] or variability or multifunctional equipment features [40–42]. The expandability or reducibility of a hospital due to growing or shrinking space requirements [12,14,25], the changeability of interior spaces [14,23], variable functions with a constant building structure [40,41] or modular hospital structures [12,37] are mostly equated with flexibility, but are to be understood far more profoundly according to the understanding of factory planning and are thus rather to be assigned to transformability. The flexibility in factory planning is described as the ability to react to foreseen changes within a defined area. Transformability, on the other hand, goes beyond flexibility and involves multidimensional changes and multiple areas [15].

In the following the general approach bridge the research gap explained above is described in Section 2. Further, Section 3 examines the prerequisites for transferring factory planning methods to hospital systems, before the paper concludes in Section 4.

2. General approach to close the research gap

Factories are also faced with the challenge of designing their processes as efficiently as possible and ensuring sustainability in a turbulent environment. Flow orientation and transformability are therefore well-known target areas in factory planning. They are already successfully taken into account within the framework of a structured procedure for the goal-oriented design of a factory's objects [15]. The factory planning tasks required for this purpose, which need expert or experiential knowledge [43], are supported by a comprehensive set of methods and design principles [15]. These are to be taken up within the framework of planning support for hospitals and adapted to the hospital system.

Factory planning distinguishes between **rough** and **detailed planning**. In the **rough planning** phase, factory planners create a feasible factory concept that fulfils the defined factory goals in the best possible way [44]. In order to achieve efficient processes, it is necessary to realise a flow orientation in the layout [45]. To support the hospital planning decisions, guiding principles are to be developed regarding centralisation or decentralisation as well as the arrangement of functional areas and rooms in accordance with the flow orientation. In **detailed planning**, the factory planners work out the factory concept and describe it in detail [44]. In this course, the factory objects, which are the material or immaterial components of a factory that it is built from [46], are designed to be transformable [47,48]. This way, factories can adequately encounter the change drivers from their turbulent environment [16,48,49]. To adapt hospitals to the change drivers that affect them, hospital planning is also to be supported with guiding principles to increase the transformability. The procedure to develop the guiding principles is illustrated in Figure 2 and described below.

Since hospitals, like factories, are to be understood as a long-lasting, complex and socio-technical systems that are subject to permanent adaptation [6], an analysis of the entire hospital as a single object of investigation is not expedient [46]. Factory planners therefore subdivide the factory into individual factory objects [47], so that these can be designed in a flow orientated and transformable manner [15].

In order to define valid guiding principles for hospital systems, the **first step** is to develop a description model of the hospital system. Analogous to factory planning, the hospital objects for later design must be identified and assigned to the previously defined system levels and the design fields [15]. In the **second step,** the flow relationships must be identified to be able to make statements about flow orientation in hospitals. For this purpose, the various flows resulting from the process flows must be combined into a holistic quantity structure and presented in a generally valid flow diagram. The service network must be depicted holistically so that all flows of patients, staff and material along the primary and secondary processes are shown. In the **third step**, target structures are to be developed based on the factory structures [15] and divided into functional areas or departments that define the process flow. With the help of process flow models and target structures, flow-oriented function diagrams are drawn up as in factory planning [15]. These serve as the basis for creating the ideal arrangement variants, from which guiding principles for the ideal arrangement of hospital areas are then to be derived. Once the guiding principles have been drawn up, the rough planning is complete. Subsequently, the guiding principles for transformability are to be worked out within the framework of detailed planning. In the **fourth step**, a maturity model must first be developed to assess the transformation potential of the individual hospital objects. In the process, individual transformation enablers for hospital systems are to be taken into account in accordance to factory planning [15], and then, characteristics of the individual factory objects that positively or negatively influence the transformation potential are to be identified. The characteristics serve as the basis for a utility analysis to assess the transformability of the whole hospital. In the **fifth step**, the required degree of transformability for the individual hospital objects must initially be determined in order to derive guiding principles for the development of space, technology and organisational concepts that meet the requirements. To this end, megatrends in the hospital environment must be identified and the resulting change drivers for hospitals and their effects on the functional areas must be derived. Subsequently, possible planning variants of the hospital objects are to be developed, from which guiding principles are to be derived depending on the necessary degree of transformability. Finally, the established guiding principles for flow orientation and transformability are to be evaluated with the help of case studies and the active involvement of hospital practitioners such as managers or physicians in workshops.

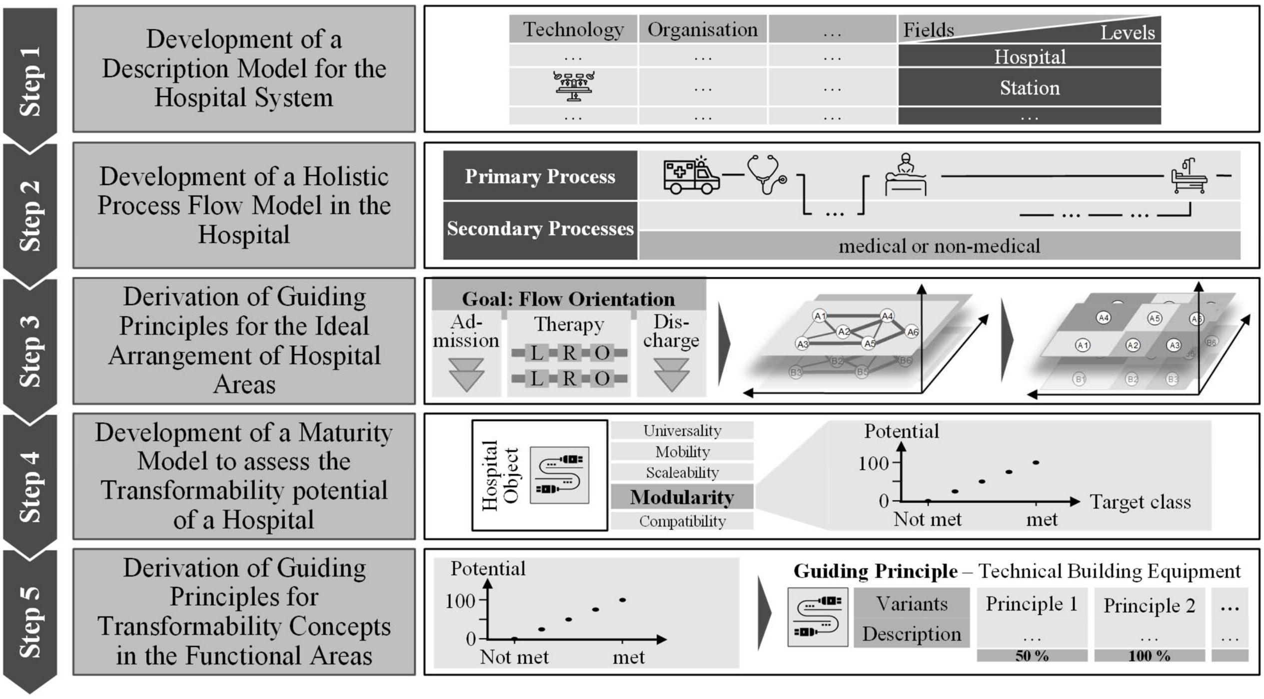


Figure 2: Procedure to develop the Guiding Principles for Flow Orientation and Transformability

3. Prerequisites for transferring factory planning approaches to hospital systems

In order to cope with external influences of cost pressure and change drivers, hospital systems are to be designed flow orientated and transformable. As described before, the approach is based on the methodical approaches from factory planning. In order to successfully transfer the described project to the hospital system, prerequisites must be met in the hospital, just as they are in factories. To design the hospital in a targeted manner, hospital objects must be identifiable (prerequisite 1). To realise flow orientation, flow relationships must exist in hospitals (prerequisite 2). In the following, the necessary similarities between the hospital and factory systems are shown and based on this, the fulfilment of the prerequisites is derived.

3.1 Hospital objects

As summarised in Figure 3, there are similarities between the fields and levels of both systems. As explained in more detail below, objects can be identified via the levels of the hospital and sorted into the matrix of fields and levels. The basic possibility of identifying hospital objects, as prerequisite 1, is thus fulfilled.

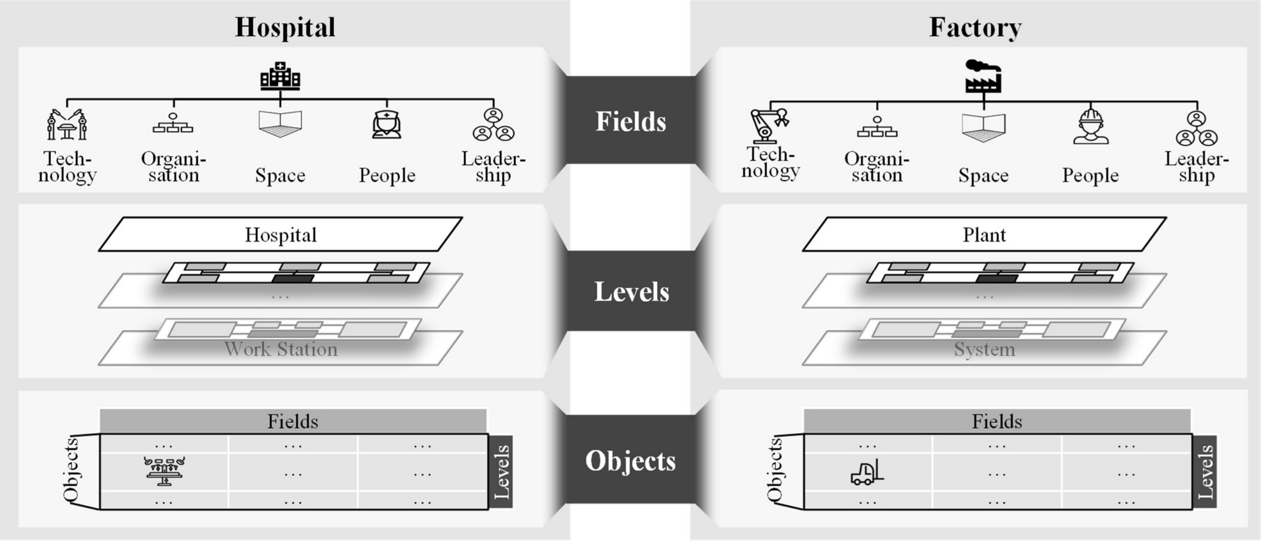


Figure 3: Similarities between Factories and Hospitals to derive Objects in Hospital Systems

Factories are divided into five factory fields of technology, organisation, space, people and leadership [46]. According to their definition, a division into these five fields can also be applied to hospital systems. Technical equipment and infrastructure are necessary for the provision of services in hospitals [6]. The objects used in hospitals differ from those used in factories, but can also be structured in production, storage, or transport. Like factories [46], hospitals have structural and procedural organisations that influence the quality of hospital services [4]. The areas in hospitals are clearly defined and structured for the building by standards and specifications according to the DIN 277 [50] and aspects such as the plot of land or the layout are also taken into account [12]. In addition, hospital employees are subject to hierarchical management [4].

The factory is hierarchically divided into five different factory levels from the plant level to the factory level, areas, sub-areas and to the work station, considering that every level encompasses all of its subordinate levels [46]. In analogy, hierarchical levels can be identified and structured for hospitals, such as from the speciality level to the work station [6]. Referring to the example that a plant may include several factory halls at the factory level, a large hospital may have several specialities at one location. These specialities are divided into various areas, e.g., surgical areas or wards, which in turn are divided into different work stations.

In the context of factory planning, material or immaterial resources are named factory objects and are classified in the matrix of factory fields and levels [46]. Different resources can also be allocated to the individual levels in the hospital and represent material or immaterial components, such as the allocation of the supply structure to the hospital level or of medications and other materials to the workplace level [6]. A definition of hospital objects and their classification into hospital fields and levels are consequently possible.

3.2 Flow relationships in the hospital system

As explained in more detail below and summarised in Figure 4, there are similarities between the products, the main processes, and the workflows of hospital and factory systems. This results in flow relationships; prerequisite 2 is therefore also fulfilled.

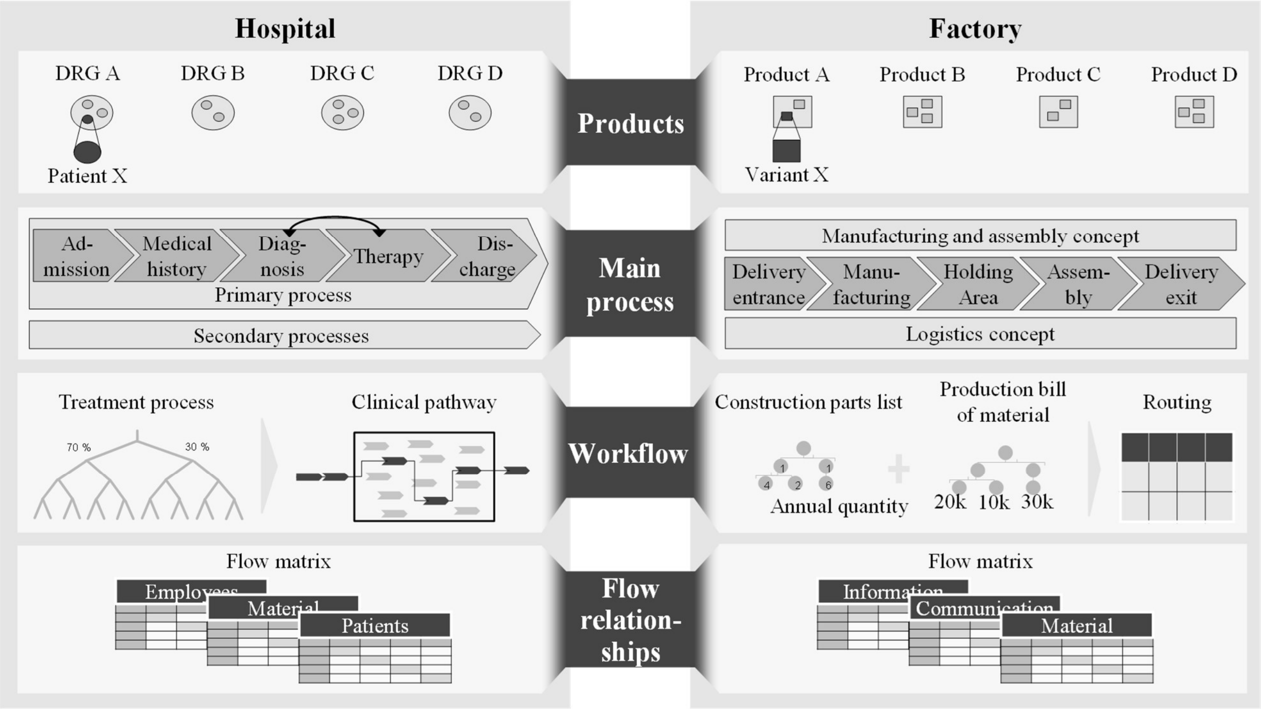


Figure 4: Similarities between Factories and Hospitals to derive Flow Relationships in Hospital Systems

Products are understood as output of a manufacturing company, which is made available to consumers for purchase [51]. The product of a hospital is the ready-treated patient. For this purpose, services are provided, such as nursing care, an examination, or an operation. In order to be able to offer the services, physical commodities or consumables are often necessary, such as medicines, X-rays or surgical instruments. The treatment of a patient depends on diagnosis, procedures, age, sex and possible complications or comorbidities, in summary DRGs. This leads to a high variance of products and to the assumption that hospitals can be understood as multi-product factories [52]. If the products of factories are differentiated into variants per product group [15], analogously the products in hospitals can be differentiated by patients per DRG.

Hospital processes required to deliver services can be divided into primary and secondary processes. Patient treatment represents the primary process of a hospital [12]. This comprises all tasks of anamnesis, diagnostics, and therapy from admission to discharge, whereby iterations between diagnosis and therapy are possible. Secondary processes support the primary process by providing the services necessary for treatment [53]. Since the necessary processes in factory systems are also subdivided, there is a further analogy between the two systems. In factories, there are the core processes for service provision, which include storage, manufacturing, and assembly from arrival to departure of goods. These are assisted by support processes [15].

Hospital and factory systems have defined workflows. For factories, this results from the work plan, which is derived from the construction and production bill of materials [15]. Workflows in hospitals are determined by clinical pathways. Clinical pathways are developed for disease patterns and group individual cases into a homogeneous group based on certain indicators or symptoms [4,54]. Starting from a diagnosis, the clinical pathway represents a consensus across professional groups and institutions for the succeeding sequence of diagnostic and therapeutic measures within the framework of inpatient treatment [20,54].

For the various diagnostic and therapeutic measures, patients pass through different clinical departments of the hospital for treatment [55]. The combination of primary and secondary processes results in a network of services in the hospital consisting of internal customer-supplier relationships [56], similar to factories [57]. This leads to a multitude of flow relationships, for example, of patients, employees, or material. In case of availability of necessary flow information, the relationships can be represented in a structured and clear manner in matrices, as they are also used in factory planning.

3.3 Differences between hospital and factory systems

Of course, hospital systems are not comparable with factory systems in all aspects. Hospitals have routine patient treatments, but also a large proportion of emergency cases. The occurrences of emergency cases is normally taken into account, however the individual and highly diverging treatments can only be determined after examining the patients. In addition, in large incidents, it is difficult to plan not only the activities to be carried out but also the number of treatments [58]. This short-term order planning is unusual for manufacturing companies; they plan in advance and adjust their factory operations accordingly. Although the emergency degree is not comparable with those in hospitals, and they do not deal with human lives, manufacturing companies in the field of maintenance, repair and overhaul are familiar with these challenges. Before they take action, similar to patients in hospitals, the product must be assessed to determine the next course of action [59,60]. In general, this fact plays a subsidiary role in transferability of factory planning methods to hospital systems, as long as the individual characteristics of hospitals are taken into account accordingly. For flow orientation, flow relationships must consider all activities and procedures, routine and emergency treatments, as well as their respective frequencies (prerequisite 1). Further, hospital objects must ensure that all possible objects from activities and processes of emergency treatment are included (prerequisite 2).

4. Summary and outlook

External influences such as cost pressure and change drivers require flow orientation and transformability of hospital systems, with internal characteristics counteracting this. Demands for flow orientation and transformability already exist for factories. Over time, factory planning has developed successful methods with which systems can be designed accordingly. Since these have been lacking for hospital systems so far, this paper presents a procedure with which the factory planning methods for flow orientation and transformability can be transferred to hospital systems. However, a condition for the successful implementation of flow orientation and transformability in hospital systems is the similarity between the two systems. Hospital objects and flow relationships between structural elements must exist in hospitals. These prerequisites were also examined, leading to the conclusion that hospitals and factories are comparable. Like factories, hospitals can be divided into levels and fields from which hospital objects can be derived. Since both systems have products and a superordinate main process, flow relationships between the structural elements can also be developed. On this basis, it is possible to describe the hospital system in a structured way analogous to factories.

Having examined the similarities and prerequisites of both hospitals and factory systems, the next step is to develop the guiding principles using the approach mentioned. The to-be-developed guiding principles should facilitate the practical and implementable transfer of factory planning methods to hospital systems.

Acknowledgements

The authors kindly thank the German Research Foundation (DFG) for funding the research project "Entwicklung von fabrikplanerischen Leitprinzipien zur Planung von flussorientierten Krankenhäusern mit wandlungsfähigen Raum-, Technik- und Organisationskonzepten (MedFAP)" (project number 456684366).

References

[1] BGBl. I. Gesetz zur Reform der gesetzlichen Krankenversicherung ab dem Jahr 2000.

[2] Walther, M., 2005. Logistik als Rationalisierungsinstrument und strategischer Wettbewerbsfaktor in einem dynamischen Marktumfeld: Auf der Suche nach operativer Exzellenz im Krankenhaus. Dissertation. Accessed 4 February 2020.

[3] Augurzky, B., Krolop, S., Liehr-Griem, A., Schmidt, C.M., Terkatz, S., 2004. Das Krankenhaus, Basel II und der Investitionsstau. RWI, Essen.

[4] Salfeld, R., Hehner, S., Wichels, R., 2009. Modernes Krankenhausmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Blöchle, D., 2014. Evaluation von Krankenhauslayouts unter Zuhilfenahme der Fuzzy-Logik: Dissertation. Accessed 4 February 2020.

[6] Fuchs, D., 2013. Entwicklung eines Vorgehensmodells zur prozessorientierten Planung und Gestaltung komplexer Arbeitssysteme (Am Beispiel Krankenhaus): Dissertation. Accessed 4 February 2020.

[7] Reuter, C., 2011. Modellierung und dynamische Adaption klinischer Pfade. Dissertation, Dortmund.

[8] Statistische Ämter des Bundes und der Länder. Demografischer Wandel in Deutschland, Heft 2, Ausgabe 2010.

[9] Loos, S., Albrecht, M., Zich, K., 2019. Zukunftsfähige Krankenhausversorgung: Simulation und Analyse einer Neustrukturierung der Krankenhausversorgung am Beispiel einer Versorgungsregion in Nordrhein-Westfalen. BertelsmannStiftung.

[10]Moesta, K.T., Schneider, J., 2020. Organisationsumbau in Corona-Zeiten : Pandemie als Katalysator. kma 25 (09), 42–44.

[11]Behar, B.I., Guth, C., Salfeld, R., 2018. Modernes Krankenhausmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg.

[12]Heinen, P., 2004. Wechselwirkungen zwischen medizinischen Prozessen und baulichen Strukturen im Krankenhausbau: unter besonderer Berücksichtigung der Entwicklung des Klinikums der Universität zu Köln. Dissertation, Köln.

[13]Braun von Reinersdorff, A., 2007. Strategische Krankenhausführung: Vom Lean Management zum Balanced Hospital Management, 2., unveränd. Aufl. ed. Huber, Hans, Bern.

[14]Wischer, R., Riethmüller, H.-U., Daschner, F., 2007. Zukunftsoffenes Krankenhaus: Fakten, Leitlinien, Bausteine. Springer, Wien.

[15]Wiendahl, H.-P., Nyhuis, P., Reichardt, J., 2014. Handbuch Fabrikplanung: Konzept, Gestaltung und Umsetzung wandlungsfähiger Produktionsstätten. Sytematik der Veränderungsfähigkeit, 2nd ed. Carl Hanser Verlag.

[16]Klemke, T., 2013. Planung der systematischen Wandlungsfähigkeit von Fabriken. Dissertation, Garbsen.

[17]Helmig, B., Tscheulin, D.K. (Eds.), 1997. „State-of-the-Art“ in der betriebswirtschaftlichen Forschung zum Krankenhausmanagement.

[18]Ekkernkamp, A., Debatin, J.F., Schulte, B., Tecklenburg, A., 2015. Krankenhausmanagement: Strategien, Konzepte, Methoden, 1. Aufl. ed. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH amp; Co. KG, s.l.

[19]Schlüchtermann, J., 2016. Betriebswirtschaft und Management im Krankenhaus: Grundlagen und Praxis, 1. Aufl. ed. MWV Medizinisch Wissenschaftliche Verlagsgesellschaft mbH amp; Co. KG, s.l.

[20]Koitka, C., 2010. Implementierung und Wirksamkeit Klinischer Behandlungspfade: - Eine systematische Literaturanalyse -.

[21]Wandschneider, W., Preiss, P., 2013. Kostendämpfung in der Herzchirurgie auf der Grundlage Klinischer Pfade: Stand und Perspektiven der Vorgehensweise in der Herzchirurgie des LKH Klagenfurt.

[22]Lux, T., Raphael, H., 2009. Prozessorientierte Krankenhausinformationssysteme. HMD Praxis der Wirtschaftsinformatik (46), 70–78.

[23]Arnolds, I.V., Nickel, S., 2013. Multi-period layout planning for hospital wards. Socio-economic planning sciences.

[24]Assem, M., Ouda, B.K., Wahed, M.A., 2012. Improving operating theatre design using facilities layout planning. Cairo International Biomedical Engineering Conference (CIBEC), 109–113.

[25]Chraibi, A., Kharraja, S., Osman, I.H., Elbeqqali, O., 2015. Optimization of dynamic operating theatre facility layout, in: Framinan, J.M. (Ed.), Proceedings of 2015 International Conference on Industrial Engineering and Systems Management (IESM). IEEE, Piscataway, NJ, pp. 262–271.

[26]Helber, S., Böhme, D., Oucherif, F., Lagershausen, S., Kasper, S., 2016. A hierarchical facility layout planning approach for large and complex hospitals. Flex Serv Manuf J 28 (1-2), 5–29.

[27]Boucherie, R.J., Hans, E.W., Hartmann, T., 2012. Health care logistics and space: Accounting for the physical build environ-ment. Winter Simulation Conference, IEEE, 696–703.

[28]Gibson, I.W., Lease, B.L., 2007. An approach to hospital planning and design using discrete event simulation. Winter Simulation Conference, IEEE, 1501–1509.

[29]Hanne, T., Melo, T., Nickel, S., 2009. Bringing Robustness to Patient Flow Management Through Optimized Patient Transports in Hospitals. Journal on Applied Analytics (39), 241–255.

[30]Nassar, K., 2010. A model for assessing occupant flow in building spaces. Automation in Construction (19), 1027–1036.

[31]Thorwarth, M., Arisha, A., 2012. A simulation-based decision support system to model complex demand driven healthcare facilities. Winter Simulation Conference, IEEE.

[32]Vos, L., Groothuis, S., van Merode, G.G., 2007. Evaluating hospital design from an operations management perspective. Health care management science 10 (4), 357–364.

[33]Acar, Y., Kadipasaoglu, S.N., Day, J.M., 2009. Incorporating uncertainty in optimal decision making: Inte-grating mixed integer programming and simulation to solve combinatorial problems. Computers & Industrial Engineering (56), 106–112.

[34]Arnolds, I.V., Gartner, D., 2018. Improving hospital layout planning through clinical pathway mining. Ann Oper Res 263 (1-2), 453–477.

[35]Baumgart, A., Denz, C., Bender, H.-J., Schleppers, A., 2009. How work context affects operating room processes: using data mining and computer simulation to analyze facility and process design. Quality Management in Healthcare (18), 305–314.

[36]Lee, H.-Y., Yang, I.-T., Lin, Y.-C., 2012. Laying out the occupant flows in public buildings for operating efficiency. Building and Environment (51), 231–242.

[37]Vera, A., Foit, K., 2012. Modulare Krankenhausorganisation und Effizienz. Zeitschrift für Betriebswirtschaft 75 (4), 357–382.

[38]Bayer, S., Köberle-Gaiser, M., Barlow, J., 2007. Planning for adaptability in healthcare infrastructure, London, 11 pp.

[39]Neufville, R. de, Lee, Y.S., Scholtes, S., 2008. Flexibiltiy in Hospital Infrastructure Design.

[40]Kyrö, R., Peltokorpi, A., Luoma-Halkola, L., 2019. Connecting adaptability strategies to building system lifecycles in hospital retrofits. Eng, Const and Arch Man 26 (4), 633–647.

[41]Neufert, E., Kister, J., Lohmann, M., Merkel, P., Brockhaus, M., 2019. Bauentwurfslehre: Grundlagen, Normen, Vorschriften über Anlage, Bau, Gestaltung, Raumbedarf, Raumbeziehungen, Maße für Gebäude, Räume, Einrichtungen, Geräte mit dem Menschen als Maß und Ziel : Handbuch für den Baufachmann, Bauherrn, Lehrenden und Lernenden, 42., überarbeitete und aktualisierte Auflage ed. Springer Vieweg, Wiesbaden, XIV, 606 Seiten.

[42]Pati, D., Harvey, T., Cason, C., 2008. Inpatient Unit Flexibility. Environment and Behavior 40 (2), 205–232.

[43]Grundig, C.-G., 2018. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 6., neu bearbeitete Auflage ed. Hanser, München.

[44]Verein Deutscher Ingenieure, 2011. Fabrikplanung Planungsvorgehen: VDI 5200.

[45]Schenk, M., Wirth, S., Müller, E., 2014. Fabrikplanung und Fabrikbetrieb. Springer Berlin Heidelberg, Berlin, Heidelberg.

[46]Heger, C.L., 2007. Wandlungspotenzialmerkmale: Bewertung der Wandlungsfähigkeit von Fabrikobjekten. Dissertation, Garbsen.

[47]Nyhuis, P., Kolakowski, M., Heger, C.L., 2005. Evaluation of Factory Transformability. 3rd International CIRP Conference on Reconfigurable Manufacturing.

[48]Wiendahl, H.-P., ElMaraghy, H., Nyhuis, P., Zäh, M., 2007. Changeable Manufacturing - Classification, Design and Operation. CIRP Annals (56), 783–809.

[49]Westkämper, E., 2002. Wandlungsfähigkeit - Herausforderung und Lösungen im turbulenten Umfeld. Sonderforschungsbereich Wandlungsfähige Unter-nehmensstrukturen für die Variantenreiche Serienproduktion (Forschungskolloquium).

[50]DIN Deutsches Institut für Normung e. V. Grundflächen und Rauminhalte im Bauwesen: Teil 1: Hochbau. Accessed 21 April 2020.

[51]Chase, R.B., Aquilano, N.J., 1981. Production and operations management: A life cycle approach, 3. ed. ed. Irwin, Homewood, Ill.

[52]Fetter, Robert B., Freeman, J. L., 1986. Diagnosis Related Groups: Product Line Management within Hospitals. The Academy of Management review (11), 41–54.

[53]Lennerts, K., 2009. Facility management of hospitals, in: Rechel, R., Wright, S., Dowdeswell, B., McKee, M. (Eds.), Investing in hospitals of the future, Copenhagen, pp. 167–186.

[54]Roeder, N., Küttner, T., Bergmann, K.O., Norbert Roeder, Tina Küttner - Deutscher Ärzte-Verlag (Eds.), 2007. Klinische Behandlungspfade: Mit Standards erfolgreicher arbeiten ; mit 4 Tabellen. Dt. Ärzte-Verl., Köln.

[55]Marsolek, I., 2003. Entwicklung einer arbeitswissenschaftlichen Methodik zur Analyse und Optimierung von komplexen Prozessflüssen im Arbeitssystem Krankenhaus.

[56]Zapp, W. (Ed.), 2010. Prozessgestaltung in Gesundheitseinrichtungen: Von der Analyse zum Controlling, 2., vollst. überarb. und erw. Aufl. ed. Economica-Verl., Heidelberg, XV.

[57]Teubner, R.A., 1999. Organization and Information Planning, in: Teubner, R.A. (Ed.), Organisations- und Informationssystemgestaltung. Deutscher Universitätsverlag, Wiesbaden, pp. 196–240.

[58]Wurmb, T., Rechenbach, P., Scholtes, K., 2017. Alarm- und Einsatzplanung an Krankenhäusern: Das konsequenzbasierte Modell. Medizinische Klinik, Intensivmedizin und Notfallmedizin 112 (7), 618–621.

[59]Eickemeyer, S.C., Mische, V., Goßmann, D., 2015. Regenerationsplanung mit belastbarem Prognosewert. Zeitschrift für wirtschaftlichen Fabrikbetrieb 110 (11), 706–709.

[60]Koppold, N., Lödding, H., 2015. Bestimmung der Arbeitsumfänge in der Instandhaltungsproduktion. Zeitschrift für wirtschaftlichen Fabrikbetrieb 110 (12), 795–798.

Biography

Lena Wecken, M.Sc. (*1995) studied industrial engineering with a focus on production technology at Leibniz University Hannover and Athens University of Economics and Business. She works as a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2021.

Dr.-Ing Shiva Faeghi (*1986) is a post-doctoral researcher at the Institute of Technology and Management in Construction at Karlsruhe Institute of Technology. Her main research interests are operations research in healthcare as well as simulation and optimization of healthcare processes. She pursues interdisciplinary research areas to address research-practice gaps in hospital and healthcare management.

Lennart Hingst, M.Sc. (*1989) studied technical logistics at the University of Duisburg-Essen. After his studies Lennart Hingst worked in the industry for years. Now he works as a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover since 2019.

Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research associate at IFA. After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003.

Prof. Dr.-Ing. Kunibert Lennerts (*1963) is holding the professorship for Facility Management at Karlsruhe Institute of Technology since 2000 and was appointed to the first university professorship in Facility Management in Germany. His main research areas are facility management and sustainability with special focus on huge real estate portfolios and hospitals.

3rd Conference on Production Systems and Logistics

Knowledge Graphs For Data And Knowledge Management In Cyber-Physical Production Systems

Beatriz Bretones Cassoli[1], Nicolas Jourdan[1], Joachim Metternich[1]
[1]TU Darmstadt, Institute of Production Management, Technology and Machine Tools (PTW), Darmstadt, Germany

Abstract

Cyber-physical production systems are constituted of various sub-systems in a production environment, from machines to logistics networks, that are connected and exchange data in real-time. Every sub-system consumes and generates data. This data has the potential to support decision making and optimization of production processes. To extract valuable information from this data, however, different data sources must be consolidated and analyzed. A Knowledge Graph (KG), also known as a semantic network, represents a net of real-world entities, i.e., machines, sensors, processes, or concepts, and illustrates their relationship. KG allows us to encode the knowledge and data context into a human interpretable form and is amenable to automated analysis and inference. This paper presents the potential of KG in manufacturing and proposes a framework for its implementation. The proposed framework should assist practitioners in integrating raw data from multiple data sources in production, developing a suitable data model, creating the knowledge graph, and using it in a graph application. Although the framework is applicable for different purposes, this work illustrates its use for supporting the quality assessment of products in a discrete manufacturing production line.

Keywords

Knowledge Graph; Cyber-Physical Production Systems; Data and Knowledge Management.

1. Introduction

The concept of Cyber-Physical Production Systems (CPPS) applies to autonomous and cooperative sub-systems (e.g. machines, sensors, actuators) that are connected and exchange data across all levels of production [1]. The data generated by CPPS, particularly from Internet of Things (IoT) devices, can be used to monitor manufacturing processes online and support decision-making. A challenge inherent to CPPS is managing this data and extracting useful information which humans can use to improve productivity and prognosis [2]. Especially for applications that use Artificial Intelligence (AI) technologies, the data delivery itself and how this data is provided (in which quality, quantity, and context) to the data analytics applications are decisive for its use [3]. Contextual information in particular allows AI applications to deal with ambiguity, thus improving predictions and their capacity of aiding the decision-making process [4]. Semantic technologies, such as Knowledge Graphs (KG), which capture context and domain knowledge information, can contribute to the explainability and acceptance of AI for realizing flexible manufacturing systems [5]. Early adopters in the manufacturing industry have been using (industrial) KG to support the integration of various data sources and to enable inference and machine processing by providing a formal semantic representation of manufacturing domain knowledge [6]. However, the use of KG in the manufacturing industry is not as disseminated as in other sectors such as finance and biomedical, having mostly prototypical implementations [7]. Guidelines and frameworks for its implementation in the

DOI: https://doi.org/10.15488/12180

manufacturing industry were not found in the current literature. This paper thus contextualizes KG, provides a novel framework and a practical example of developing an industrial KG for a CPPS. This publication aims at providing guidelines for practitioners and researchers who are not yet familiar with the usage and capabilities of KG. The goal is to break the initial barrier to understanding how to develop a KG, especially for manufacturing data and facilitating the implementation of this semantic technology.

This paper is structured as follows; section 2 presents the theoretical background of KG, graph databases and ontologies. The goal is to situate KG in the broader spectrum of AI technologies. Section 3 introduces the methodology and the proposed framework. Finally, section 4 concretizes the framework by providing a step-by-step development of a KG application having as an example an openly accessible dataset provided by Bosch of four of its assembly lines. The section concludes with an outlook of how the developed KG can be used as part of a Machine Learning (ML) pipeline to predict part quality. The fifth and last section provides future work directions.

2. Theoretical Background

A Knowledge Graph (KG) is a collection of nodes (for each entity), and edges, representing the relationships which connect and relate these entities to the world [8]. Figure 1 provides an example of a basic graph. In this graph, three entities are shown (a person, a station, and a machine), together with the directed relationships that connect them (a person works in a station, and a station has a machine as resource).

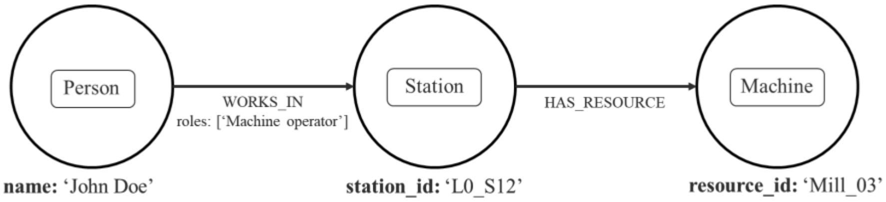


Figure 1: Nodes and relationships

KG are particularly useful because they add a layer of metadata to the data, providing context to it and defining rules for its structure and interpretation [9]. KG can, therefore, represent complex relationships in a domain in both machine friendly and human-readable forms [9]. Graph databases are databases developed specifically to store graph-shaped data and are designed to take advantage of the data connectivity to extract information and generate insights [10]. The uses of KG and of graph databases can typically be divided into three areas: graph statistics (statistic measures about the graph), graph analytics (analysis of the graph data to answer specific questions through queries or graph algorithms), and graph-enhanced ML and AI (application of graph data and analytics results to create or select features for ML models) [11]. In the context of graph-enhanced ML, KG embeddings are techniques used to insert entities and relationships from a KG into continuous vector spaces to allow further mathematical manipulation (e.g. use as input data for ML models) while still preserving the KG structure [12]. KG embeddings are being used in ML pipelines to improve the learning process through the integration of semantic rich features [4].

An ontology represents a domain, its objects, and the formalized relationships between them in a declarative formalism [13]. It is a human-readable text that describes entities in a domain and contains formal axioms (propositions) constraining the interpretation and use of these terms [13]. An ontology is used to share a common understanding of the information structure between people and machines, formalize and thus allow the reuse and analysis/reasoning with domain knowledge, make domain assumptions explicit, and, finally, separate domain knowledge from the operational knowledge [14]. KG and ontologies are part of the broader field of Knowledge Representation and Reasoning, a subfield of AI. Ontologies can be applied to define the KG's underlying structure and guarantee interoperability with other systems and applications from the same domain [15].

Graph technologies are particularly suitable if the relationships within the data are central for analytics and reasoning purposes. Although relational systems (also called SQL databases) are well documented and easier for developers, retrieving relationships out of relational systems is considerably more complex and less efficient than graph systems [16].

Recent literature on graph databases deals with whether graph technology is appropriate for specific use cases. Gosnell et al. [16] propose three questions to decide whether a particular problem is suited for graph technologies:

1. Does the problem need graph data?
    - Understand the shape of the data the problem requires – when understanding the use case, is it natural to represent the data in a graph shape? What is the type of information to be extracted from the data?
2. Do relationships between entities in the data help understand and solve the problem?
    - Are the relationships between entities in the data relevant to explain and clarify the problem?
3. What is the purpose of exploring the relationships in the data?
    - Is the goal to generate a report (classical business intelligence)? Perform research on graph-structured data? Provide a service to an end-user (retrieval)? Use as part of an ML pipeline?

Nowadays, KG are being applied in various industries in application areas such as question answering, recommender systems, information retrieval, and feature engineering [17]. KG are also being adopted in the manufacturing industry, mostly in the form of prototypes. Zhou et al. use a KG-based approach for resource allocation in a discrete manufacturing workshop [18]. To achieve a self-configurable manufacturing process, Zheng et al. introduce a KG-based multi-agent reinforcement learning method [19]. Zhao et al. developed a service platform based on KG for resource allocation and scheduling of manufacturing orders [20]. Additional use cases include the use of KG for creating the digital twin of a building used in applications for building management and services, risk management in engineering projects, process monitoring, and machine service operations planning [6, 21].

3. Methodology and proposed framework

This section introduces a framework for developing and using KG. The framework is based on the literature presented in the theoretical background section, mostly focused on the use of graph databases. Mainly the work from Robinson et al. [8], Barrasa et al. [9], Gosnell et al. [16], Sequeda and Lassila [22], combined with practical implementation experience from the authors, are translated into a general guideline for the adoption of KG. Figure 2 presents the framework consisting of five phases. The three questions proposed by Gosnell et al. [16] and introduced in section 2 are to be answered before the first phase of the framework. Next, each phase will be explained from a theoretical point of view.

1. Problem understanding and identification of data sources. As a natural first step, the problem understanding aims at having a clear goal for the KG application. In this step, the purpose of exploring the relationships in the data should be differentiated between generating a business report (calculating key performance indicators), providing a service to an end-user (retrieval system), or being used as part of an ML pipeline. The identification of data sources and their relationships aids the data modelling. We suggest conducting a workshop with stakeholders for the problem understanding step.

2. Data modelling. The data model represents the domain's entities and relationships and establishes the graph structure for receiving data in the subsequent steps. This step is paramount and should be guided by the goal defined in the problem understanding. When querying a graph, we talk about "traversing", which means going from one point (or node) of the graph to another. So how the entities and their relationships are represented in the data model will affect the traversals and, therefore, the efficiency of information retrieval and how the domain is represented in the graph. One of the biggest differences and advantages from graph data storage is that the conceptual data model (how the entities relate to each other in the domain, for example the one portrayed in Figure 1) translates directly to how the data is physically stored in the graph database [16]. We identify two main ways of defining the data model: through an existing or newly developed ontology stored, e.g., in a Resource Description Framework (RDF) file (**2.1**) or through graph modelling frameworks, such as the Labeled Property Graph (LPG) [8] or the Graph Schema Language (GSL) [16] (**2.2**). Independently of the chosen approach, each entity or object from the domain is assigned a specific node label. Node labels are used for entities from the same class, which will share the same type of relationships and properties. Similarly, relationships' labels define their type and properties. This step can likewise be conducted in a form of a workshop with stakeholders and using visual tools, e.g., whiteboards, to draw the entities and their relationships and derive the data model.

3. Data preparation and integration. In most cases, the data sources come from relational databases or are stored in a relational data format, such as tables in CSV files or as unstructured data in NoSQL databases. The question in this task is how to prepare the data and integrate the various data sources to be loaded into the specified graph data model defined in step 2. The data preparation can be as simple as creating new CSV files that translate the relationship within the data that will be then loaded into the KG. For more complex scenarios, there are the so-called "mappings". A mapping is "a function representing the relationship from a source data model to a target data model" [22]. Mappings are usually formulated as rules (IF (condition) THEN (conclusion)) and indicate how a source (e.g., relational database) can be represented in the target (a KG). These functions or rules are applied to fill the KG with data originating from relational or NoSQL systems. At this stage, the graph database (such as Neo4j, DataStax, Amazon Neptune, and Cassandra) should be chosen, as it will influence the necessary changes in the original data for its integration.

4. Knowledge Graph creation. Once the data sources are identified and prepared for integration, the next step is to load the data into the selected graph database. As mentioned before, different procedures depending on the graph database exist for conducting this task. The basic idea is to load the entities and relationships for the data as defined in the graph data model in step two. At the end of this step, entities and relationships from the data should be represented and connected in the KG. The KG can then be used to develop the application in the fifth step.

5. Application development. The fifth and last step of the framework, application development, summarizes the efforts for creating an application that will ultimately use the KG. These uses can be assigned to three main categories: Generating business reports (**5.1**), querying, and retrieving information for an end-user (**5.2**), or as part of an ML pipeline for either producing graph features or facilitating the selection of features, e.g., for regression or classification tasks (**5.3**).

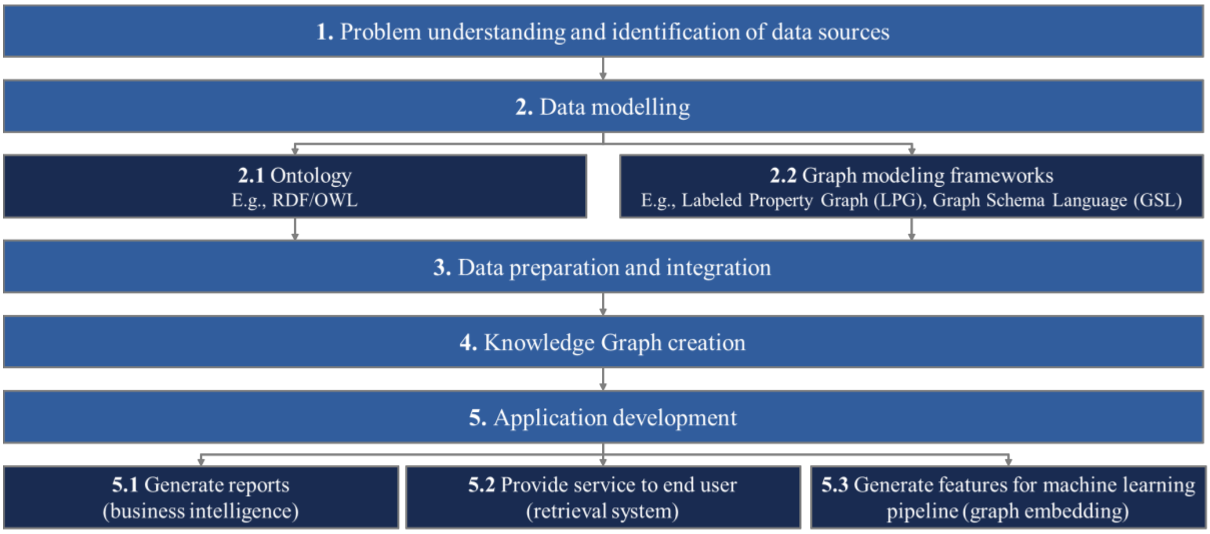


Figure 2: Proposed framework

4. Framework Application: Case Study Product 360

A well-known use case for KG in enterprises is the so-called "Customer 360" applications. Customer 360 are considered data products that consolidate and connect various data sources about the customer, creating a common approach to expose this information for use [22]. The 360-degree view of customer's data allows using the available and most relevant information about each customer to improve and customize service offerings, e.g., products recommendations [23]. Similarly, we propose a 360-degree view of the product. The idea is to gather in a KG all available and relevant information about a manufactured product to support additional analysis, e.g., for creating graph features for an ML pipeline to predict which part will fail quality control (predictive quality). In this work, steps one to four (from problem understanding to KG creation) are presented to clarify how to use the framework. The fifth step (application development) gives an outline and future work directions for the predictive quality case. For this application development, the graph database Neo4j through the Neo4j Desktop application was used.

1. Problem understanding and identification of data sources. The data used to develop the product 360 KG is from the publicly available datasets "Bosch Production Line Performance" published in Kaggle in 2016[1]. The datasets contain measurements of parts as they move through Bosch's assembly lines and are divided into three types of features plus a response feature: numerical, categorical, date (timestamps) and binary labels indicating whether the part succeeded (Response = 0) or failed (Response = 1) quality control. The datasets are further divided into training and testing sets for each feature type. The training and testing sets have 1.184.687 and 1.183.748 samples, respectively. Further, the data contains 968 numerical features, 2.140 categorical features and 1.156 date features. This results in a combined dataset of 14,3 Gigabyte. Besides the challenge of working with large datasets, the ground truth is highly imbalanced (approximately one in every 125 products are defective), and several product variants are represented in the dataset (identified by the 7.148 unique flow paths) [24]. The identified problem is then predicting which parts will fail quality control. Therefore, the data sources are the training and testing datasets: train and test numerical, train and test date, and train and test categorical datasets.

2. Data modelling. The data model should aid the task of identifying which products pass or fail the quality control. The goal is to generate a KG with all information about that product. In this case, it means representing in the graph the numerical and categorical features related to each station and line, the sequence of stations and lines, and the data features in the form of timestamps. Figure 3 shows the data model developed using the LPG framework.

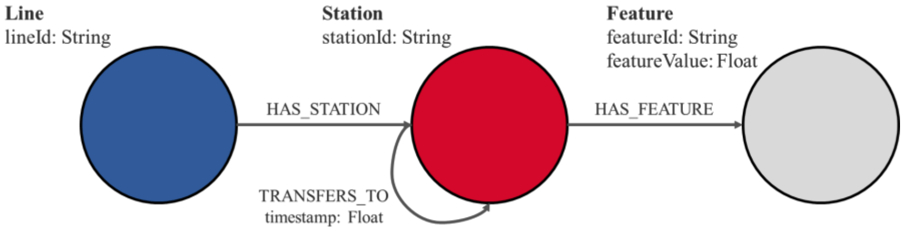


Figure 3: LPG data model for product 360 knowledge graph

In the figure, we identify three types of entities: Line, Station, and Feature. Lines are identified by the "lineId", which is saved as a String. Similarly, stations are identified by a "stationId" stored as a String. Features have two properties: the identifier "featureId" saved as String, and the "featureValue" stored as a Float and which holds the actual feature value (recordings from sensors or categorical data). The data model further defines three types of relationships: HAS_STATION, HAS_FEATURE, and TRANSFERS_TO. HAS_STATION is used to describe which stations the line comprises. HAS_FEATURE indicates in which

---

[1] Available at: https://www.kaggle.com/c/bosch-production-line-performance/data

station the feature was recorded. Finally, TRANSFERS_TO captures the sequence of stations and lines the product passed through during the assembly process. The “TRANSFERS_TO” relationship has a property called “timestamp”, which records the timestamp of the estimated time the transfer between stations occurred. This data model serves as the basis for preparing the data and loading it in the Neo4j graph database. Alternatively, an ontology can be used as the data model and imported in Neo4j using the neosemantics (n10s) plugin. Figure 4 shows an exemplary ontology created using WebProtègè[2] and exported as a turtle (.ttl) file.

Figure 4: Ontology in WebProtègè

3. Data preparation and integration. The data preparation step consists of reshaping the Bosch production line dataset to be loaded into the graph database. Table 1 presents an excerpt of the first five rows and five columns of the raw numerical train dataset. This dataset was used to capture the entities for the Line, Station and Feature classes and the “featureValue” for the numerical features.

Table 1: Excerpt of Bosch's production line numerical train dataset

Index	**Id**	**L0_S0_F0**	**L0_S0_F2**	**L0_S0_F4**	**L0_S0_F6**
0	4	0.03	-0.034	-0.197	-0.179
1	6	nan	nan	nan	nan
2	7	0.088	0.086	0.003	-0.052
3	9	-0.036	-0.064	0.294	0.33
4	11	-0.055	-0.086	0.294	0.33

Table 2 depicts the result of shaping the data for the first product in the dataset, product Id 4. For manipulating the raw data, the Python library pandas was used. The pandas data frame was saved into a CSV file which will be used in the subsequent step to create the KG. This table allows identifying all entities, their respective Ids, and the relationships “HAS_STATION” and “HAS_FEATURE”.

Table 2: Entities and feature values for product Id 4

Index	**lineId**	**stationId**	**featureId**	**featureValue**
0	L0	L0_S0	L0_S0_F0	0.03
1	L0	L0_S0	L0_S0_F2	-0.034
2	L0	L0_S0	L0_S0_F4	-0.197
3	L0	L0_S0	L0_S0_F6	-0.179
4	L0	L0_S0	L0_S0_F8	0.118

---

[2] Available at: https://webprotege.stanford.edu/

The relationship “TRANSFERS_TO” was obtained from the date train dataset. The timestamps in this dataset were used to determine the station sequence and the estimated timestamp of transfer between stations. Table 3 shows the resulting table with the station of origin (stationId_s, s for the subject) and the destination (stationId_o, o for the object), and the timestamp listed in the column “transfers_to”.

Table 3: TRASNFERS_TO relationship

Index	stationId_s	transfers_to	stationId_o
0	L0_S0	82.24	L0_S1
1	L0_S1	82.24	L0_S2
2	L0_S2	82.24	L0_S4
3	L0_S4	82.26	L0_S7
4	L0_S7	82.26	L0_S8

4. Knowledge Graph creation. The CSV files were used to create and save the data into the KG in Neo4j. Cypher commands to add the data to the graph database were written in the Neo4j Browser application. An excerpt of the commands used to create the nodes (entities) and relationships are listed in Table 4.

Table 4: Excerpt of Cypher commands for creating nodes and relationships

Command explanation | **Cypher command**

Create the nodes for lines, stations and features from product Id 4’s CSV file

```
LOAD CSV FROM 'file:///productid4.csv' AS row
WITH row[1] AS lineId, row[2] AS stationId, row[3] AS featureId, toFloat(row[4]) AS featureValue
MERGE (l:Line {lineId: lineId})
MERGE (s:Station {stationId: stationId})
MERGE (f:Feature {featureId: featureId})
  SET f.featureValue = featureValue
RETURN count(l);
```

Create relationships “HAS_FEATURE” from product Id 4’s CSV file

```
LOAD CSV WITH HEADERS FROM 'file:///productid4.csv' AS row
WITH row.lineId AS lineId, row.stationId AS stationId, row.featureId AS featureId
MATCH (l:Line {lineId: lineId})
MATCH (s:Station {stationId: stationId})
MATCH (f:Feature {featureId: featureId})
MERGE (s)-[rel:HAS_FEATURE]->(f)
RETURN count(rel);
```

Visualize “TRANSFERS_TO” relationship and numerical features

```
MATCH (s:Station)-[rel:TRANSFERS_TO]->(s:Station)
MATCH (s:Station)-[rel2:HAS_FEATURE]->(f:Feature)
RETURN s, f, rel, rel2;
```

The result from the last query can be visualized in Figure 5. Figure 6 provides a detailed view of the "TRANSFER_TO" property, indicating the transfer's timestamp to the next station occurred.

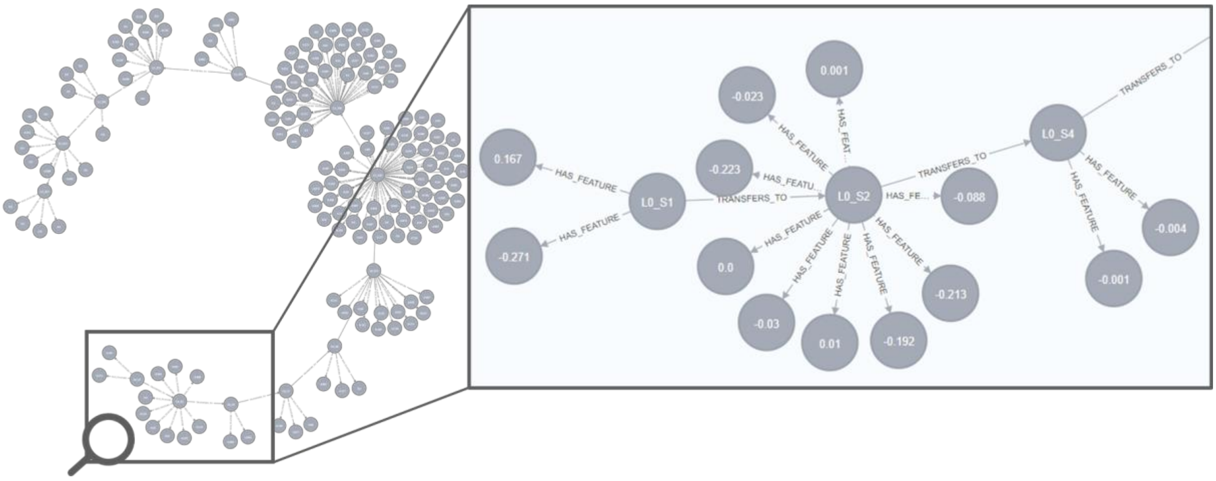
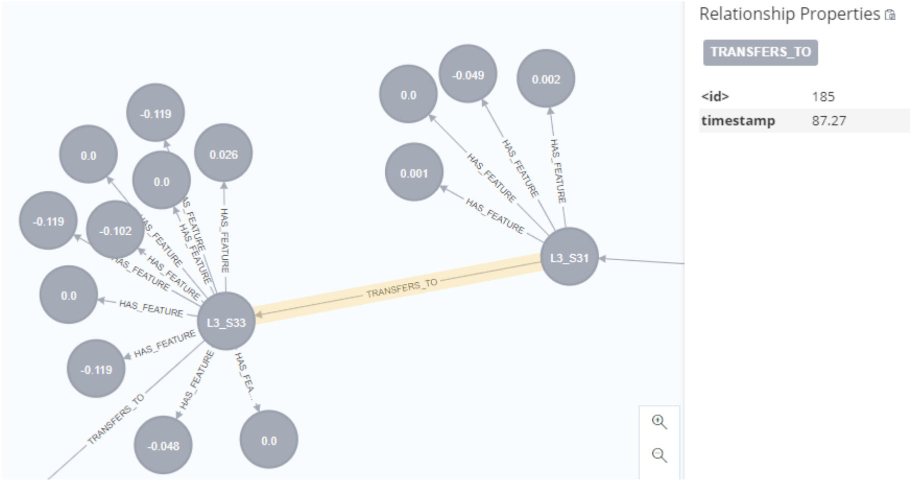


Figure 5: Knowledge Graph in Neo4j for product Id 4



Relationship Properties

TRANSFERS_TO

<id>	185
timestamp	87.27

Figure 6: Relationship property for product Id 4

The same procedure used to create the KG for product Id 4 is applied to the remaining products. The result is one KG per product.

Outlook: 5. Application development. The KG created in step four can now be used for various applications, from generating key performance indicators for each product and the assembly line and part of an ML pipeline for graph classification. Here we outline a predictive quality application currently being developed and seen as future work. The predictive quality application consists of conducting a graph classification task. First, it is necessary to generate low-dimensional vector representations of the KG created in step four through KG embeddings. For that, GraphSAGE[3] can be used for products from the same variant and, therefore, present the same stations and lines sequence. The graph embeddings are used for training a classifier which should then predict whether the assembled product has Response = 0 (succeeded the quality control) or Response = 1 (failed the quality control).

5. Conclusion

This paper presented step-by-step guidance to create a KG from relational data sources. It starts with framing the problem as a graph to develop a data model and load it into the graph database. The guideline should assist practitioners and researchers who want to use semantic technologies as a source of information. Either

---

[3] Available at: http://snap.stanford.edu/graphsage/

for calculating key performance indicators and generating business reports by ingesting data into the KG and performing graph analytics or as a retrieval system for providing a service to an end-user or as part of an ML pipeline. Future work includes providing more insights into the application development and exploring the potential to enrich feature expressiveness for ML pipelines.

Acknowledgments

The research leading to these results has received funding from the European Commission's H2020 program under grant agreement number 958357 (InterQ).

References

[1] Francalanza, E., Borg, J., Constantinescu, C., 2017. A knowledge-based tool for designing cyber physical production systems. Computers in Industry 84, 39–58.

[2] Saqlain, M., Piao, M., Shim, Y., Lee, J.Y., 2019. Framework of an IoT-based industrial data management for smart manufacturing. Journal of Sensor and Actuator Networks 8 (2), 25.

[3] Janssen, M., Brous, P., Estevez, E., Barbosa, L.S., Janowski, T., 2020. Data governance: Organizing data for trustworthy Artificial Intelligence. Government Information Quarterly 37 (3), 101493.

[4] Needham, M., Hodler, A.E., 2019. Graph Algorithms: Practical Examples in Apache Spark and Neo4j. O'Reilly Media, United States of America.

[5] Bretones Cassoli, B; Jourdan, N; Nguyen, P. H.; Sen, S.; Garcia-Ceja, E.; Metternich, J., 2021. Frameworks for data-driven quality management in cyber-physical systems for manufacturing: A systematic review, 15th CIRP Conference on Intelligent Computation in Manufacturing Engineering - CIRP ICME '21 Virtual Conference, 14-16 July 2021, Procedia CIRP, Elsevier, ISSN: 2212-8271, (in print).

[6] Hubauer, T., Lamparter, S., Haase, P., Herzig, D.M., 2018. Use Cases of the Industrial Knowledge Graph at Siemens, in: International Semantic Web Conference (P&D/Industry/BlueSky).

[7] Chen, X., Jia, S., Xiang, Y., 2020. A review: Knowledge reasoning over knowledge graph. Expert Systems with Applications 141, 112948.

[8] Robinson, I., Webber, J., Eifrem, E., 2015. Graph Databases: New Opportunities for Connected Data. O'Reilly Media, Inc., United States of America.

[9] Barrasa, J., Hodler, A.M., Webber, J., 2021. Knowledge Graphs: Data in Context for Responsive Businesses. O'Reilly Media, United States of America, 87 pp.

[10] Janev, V., Graux, D., Jabeen, H., Sallinger, E., 2020. Knowledge Graphs and Big Data Processing 12072.

[11] Hodler, A., Needham, M., 2021. Graph Data Science (GDS) For Dummies®, Neo4j Special Edition. John Wiley & Sons, Inc., United States of America.

[12] Wang, Q., Mao, Z., Wang, B., Guo, L., 2017. Knowledge Graph Embedding: A Survey of Approaches and Applications. IEEE Trans. Knowl. Data Eng. 29 (12), 2724–2743.

[13] Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowledge acquisition 5 (2), 199–220.

[14] Noy, N.F., McGuinness, D.L., 2001. Ontology development 101: A guide to creating your first ontology. Stanford knowledge systems laboratory technical report KSL-01-05.

[15] Ko, H., Witherell, P., Lu, Y., Kim, S., Rosen, D.W., 2021. Machine learning and knowledge graph based design rule construction for additive manufacturing. Additive Manufacturing 37, 101620.

[16] Gosnell, D., Broecheler, M., 2020. The Practitioner's Guide to Graph Data: Applying Graph Thinking and Graph Technologies to Solve Complex Problems. " O'Reilly Media, Inc.".

[17] Zou, X., 2020. A survey on application of knowledge graph, in: Journal of Physics: Conference Series. IOP Publishing, p. 12016.

[18] Zhou, B., Bao, J., Li, J., Lu, Y., Liu, T., Zhang, Q., 2021. A novel knowledge graph-based optimization approach for resource allocation in discrete manufacturing workshops. Robotics and Computer-Integrated Manufacturing 71, 102160.

[19] Zheng, P., Xia, L., Li, C., Li, X., Liu, B., 2021. Towards Self-X cognitive manufacturing network: An industrial knowledge graph-based multi-agent reinforcement learning approach. Journal of Manufacturing Systems 61, 16–26.

[20] Zhao, Y., Liu, Q., Xu, W., 2017. Open Industrial Knowledge Graph Development for Intelligent Manufacturing Service Matchmaking, in: 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII). 2017 International Conference on Industrial Informatics - Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), Wuhan. 02.12.2017 - 03.12.2017. IEEE, pp. 194–198.

[21] Fensel, D., Simsek, U., Angele, K. (Hg.), 2020. Knowledge Graphs: Methodology, Tools and Selected Use Cases. SpringerLink (Ed.), Austria.

[22] Sequeda, J., Lassila, O., 2021. Designing and Building Enterprise Knowledge Graphs. Synthesis Lectures on Data, Semantics, and Knowledge 11 (1), 1–165.

[23] Hristoski, I., Dimovski, T. (Hg.), 2020. Graph Database Modeling of a 360-Degree E-Customer View in B2C E-Commerce. International May Conference on Strategic Management IMCSM20. Bor, Serbia, September 25 - 27.

[24] Mangal, A., Kumar, N. (Hg.), 2016. Using big data to enhance the Bosch production line performance: A Kaggle challenge. IEEE International Conference on Big Data.

Beatriz Bretones Cassoli, M. Sc. has been a researcher and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2019. Her research interests include using semantic technologies for manufacturing use cases, particularly Knowledge Graphs.

Nicolas Jourdan, M. Sc. Has been a researcher and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2020. His research interests include the robustness analysis of Machine Learning models for manufacturing applications.

Prof. Dr.-Ing. Joachim Metternich has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2012. In addition, Prof. Metternich is the spokesman for the Center of Excellence for Medium-Sized Businesses in Darmstadt and president of the International Association of Learning Factories.

3rd Conference on Production Systems and Logistics

Development Of A Competence-based Role Model For Managers Considering Current Megatrends

Barbara Tropschuh[1], Benedict Beck[1]
[1]Technical University of Munich, Institute for Machine Tools and Industrial Management (iwb), Boltzmannstrasse 15, 85748 Garching, Germany

Abstract

Due to the digital transformation and global megatrends, the industrial environment of manufacturing companies is changing faster than ever. As a result, the tasks of managers in these companies are evolving. New management roles and competences are required to master the challenges of this industrial change and remain competitive in the global market. This paper describes the development of a competence-based role model of managers in manufacturing companies focusing on the influence of current megatrends and the associated digital transformation. The model provides an overview of future tasks and roles of managers, which are becoming increasingly important in view of current megatrends. In regard to existing role models, seven roles of modern managers are derived. These management roles were evaluated in an extensive survey and detailed with corresponding competence profiles.

Keywords

Digital Transformation; Megatrends; Competences; Management Roles;

1. Introduction

With its technological and social changes, the digital transformation significantly impacts manufacturing companies [1]. Due to current megatrends like digitalization (e.g., big data, KI, Robotik, IoT), globalization (e.g., international competition, higher customer demands), volatility (e.g., uncertainty, instability, shortened product cycles), sustainability (e.g., environmental protection, climate change, decarbonization, circular economy), and demographic change (e.g., aging, shortage of skilled workers, lifelong learning, changing values), the industrial environment is evolving [2]. These challenging environmental conditions can also be described by the acronym VUCA (volatility, uncertainty, complexity, ambiguity) [3]. In today's VUCA world, the day-to-day tasks of managers are highly complex and dynamic. Therefore, managers are constantly facing new challenges. Especially in the course of the digital transformation, the tasks of managers are changing [4]. There is no scientific consensus on how managers' tasks will change in the context of the digital transformation and global megatrends, what roles they will take on, and what competences they will need.

This paper aims to develop a competence-based role model of managers in manufacturing companies that considers the influence of current megatrends and the accompanying digital transformation. After describing the literature-based fundamentals of competence-based role models in section two, the description of the research method follows in section three. Section four includes the results of the role description and the empirical study results determining the competence requirements of the different management roles. The results will be presented in the form of competence profiles that are tailored to the corresponding role models. A discussion of the results follows before this contribution summarizes with a conclusion.

DOI: https://doi.org/10.15488/12189

2. Management role models

Based on the classic process- and task-oriented management models, the models of Koontz & O'Donnell [5] and Rühli [6] need to be mentioned. The former describes the management process and the associated tasks in five areas: planning, organization, personnel deployment, leadership, and control [5]. According to Koontz & O'Donnell [5], the central tasks of managers in these areas follow a chronological sequence. The tasks should be fully described and structured. A similar approach can be found in the Zurich management approach, according to Rühli [6], in which the four management functions, planning, decision, arrangement, and control, are distinguished. The main difference to Knootz & O'Donnell's management process is that the Zurich management model focuses less on the specific tasks of the respective manager and more on the overall management tasks of a company.

In contrast to the above-mentioned process- and task-oriented models, Mintzberg's [7] and Quinn's [8] approaches are role-based. Mintzberg identifies ten management roles, which are divided into three superordinate groups [7]: Roles in establishing and maintaining interpersonal relationships (figurehead, leader, liaison), informational roles (monitor, disseminator, spokesperson), and roles in decision-making (entrepreneur, disturbance handler, resource allocator, negotiator). The ten management roles illustrate in a practical way the broad spectrum of tasks of a manager and are described precisely [7]. At the same time, they are kept general to ensure relevance even in today's world. Quinn's competing values framework examines which roles managers assume depending on the company's strategic orientation. Based on a study of organizational effectiveness [9], Quinn [8] derives the bipolar scales of flexibility vs. stability and internal vs. external focus. In the four quadrants of the two-dimensional model, eight specific management roles are described: facilitator and mentor, innovator and broker, producer and director, coordinator, and monitor [10]. The main difference to the approach of Mintzberg [7] is that Quinn [8] explicitly names the contradictions of different roles and processes within the framework of a coherent overall model [10]. According to Quinn, effective leadership behavior exists when all roles are fulfilled [8]. This includes the role-immanent conflict that a manager sometimes has to take on opposing roles simultaneously [11].

Reinhart describes the range of management tasks based on six categories: providing orientation and goals, deciding and taking responsibility, delegating, monitoring and controlling, informing and communicating, motivating, and developing [12]. The effects of digitalization and demographic change on the above-mentioned management tasks are explicitly examined. Zeichhardt analyses the influence of digitalization on managers [13]. Here, not only the tasks of individual managers are analyzed, but the digital management tasks in the entire company are considered. These are classified based on seven digital management roles: digital figurehead, digital game changer, digital broker, agile facilitator, digital specialist, big data manager, and artificial intelligence. The management concept *Teamlead* of Graf et al. examines how managers efficiently lead teams in the context of digitalization and a VUCA environment [14]. The Teamlead model defines six system functions with 23 management tasks important for teams and leaders. The six functions are difference, resource, structure, process, reflection, and development management [14]. In the Teamlead concept, the tasks of a manager are viewed from a modern, team-based perspective.

The analyzed approaches were assessed in terms of model structuring, the considered focus area, and clustering. Regarding **model structuring**, a distinction can be made between process-oriented, task-oriented, and role-oriented models [15,16]. The former is characterized by a chronological sequence of process steps, whereas individual management tasks are unordered in a task-oriented structure. A role-oriented structure can be distinguished by summarizing various subtasks in a role description. Most of the models examined are process or task-oriented; only Mintzberg [7] and Quinn [8] focus on the individual management roles. Furthermore, the models can be differentiated based on the chosen **focus area**. In contrast to a general focus, a trend-oriented focus takes current developments and megatrends into account. Current megatrends are considered in the approach of Reinhart [12]. In the other approaches, either no recent developments or only a few selected trends are addressed. In addition, most of the mentioned approaches developed a general, exhaustive model, i.e., to represent all tasks or management roles [5,6]. Therefore, unspecific **categories and clusters** are formulated. Older models do not lose relevance because current trends hardly influence the results. However, the practical significance regarding current developments is diminished. The approach of Zeichhardt [13] describes the different management roles and tasks in a specific and concise manner. It can be stated that no model describes the roles of managers arising from current megatrends in the context of individual role descriptions. In addition, there is no competence model in which specific and prioritized competence profiles are assigned to roles arising from current trends. Figure 1 shows an overview of the evaluated role models of managers.

	model structuring			focus area		categories & clusters	
	process-oriented	task-oriented	role-oriented	general	trend-oriented	general & exhaustive	specific & concise
Koontz & O'Donell (1955)	●	○	○	●	○	●	○
Rühli (1985)	●	◐	○	●	○	●	○
Mintzberg (1973)	○	○	●	●	○	●	○
Quinn (1988)	○	○	●	●	○	●	○
Reinhart (2017)	○	●	○	○	●	●	○
Zeichhardt (2018)	○	◐	◐	○	●	○	●
Graf et al. (2020)	○	●	○	○	●	●	○

Fulfillment level: ● fulfilled ◐ half fulfilled ○ not fulfilled

Figure 1: Analysis of relevant management task and role models

3. Research method

The research method used to develop the competence-based role model is divided into two steps: In the first step, individual preliminary roles of managers were identified. These were developed based on current megatrends and managers' existing tasks and role models. The following process was used to create the preliminary roles. With the help of developed guiding questions (e.g. 'How are the potential tasks and roles of managers changing with regard to the addressed megatrends?), hypotheses were collected from the presented literature and megatrends (e.g., change of a current role due to digitalization). The hypotheses describe aspects of managers' everyday work that are becoming increasingly important due to current megatrends. Afterwards, the hypotheses were combined and clustered to derive specific management roles.

The second step represents an empirical survey among managers. The empirical-inductive procedure evaluated the preliminary roles from step one in a practical manner. For this purpose, the preliminary roles were integrated into a comprehensive survey concept. The empirical survey comprises two objectives: First, the preliminary developed roles were validated and examined regarding their relevance. Secondly, competence profiles for the respective roles were derived from an extensive collection of competences. Thereby, survey participants selected what they considered the five most important competences for each role from a collection of 40 future management competences. This competence collection was derived based on literature from Heyse & Erpenbeck [17], IFIDZ [18], Stifterverband & McKinsey [19] and Cloots [20]. 60 production-related managers from various age groups, company sizes, and management levels took part in the online survey (17 % top management, 48 % middle management, 27 % operational management, 8 % consultants).

4. Description of the results – management roles and their required competences

The preliminary roles of managers developed in step 1 were assessed by the survey participants in step 2 as part of the empirical survey. The role model contains seven roles of modern managers. According to the survey results, each management role will become more significant in the future. The seven management roles are presented in more detail below.

As a **Data Analyst**, a manager uses digital technologies to analyze large amounts of data and derives insights for management decisions. The task of a manager is to interpret the results of the data analysis and draw conclusions for their actions. Managers are thus enabled to make well-founded and data-based decisions.

As a **Change Manager**, a leader organizes change processes from an initial state to a target state. The industrial environment is constantly changing. Therefore, companies need to adapt their products or services, react to unexpected events, and manage sudden crises. Managers have the task of responsibly shaping these changes and involving employees in the process.

As a **Digital Communicator**, a manager uses digital tools to interact with people. Leadership tasks are performed digitally, and communication with employees, customers, and other stakeholders gradually shift into the digital space.

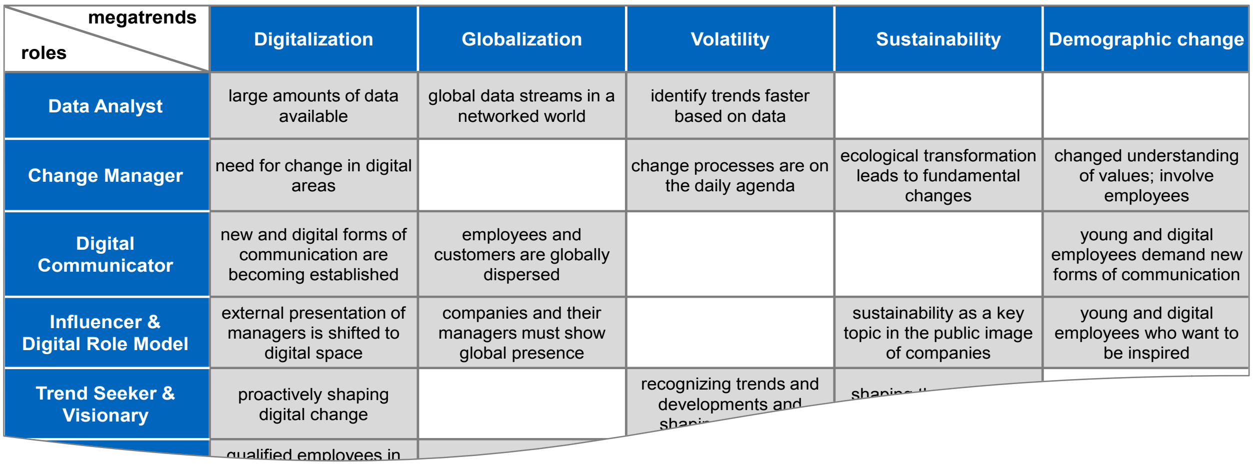
As an **Influencer & Digital Role Model**, a manager represents the company or department internally and externally. Thereby, the manager acts as a symbolic figure in the digital space and influences the opinion and behavior of people in his or her network. This includes, for example, customers, employees, investors, and journalists. By using digital services, managers multiply their reach.

As a **Trend Seeker & Visionary**, a manager engages deeply with current and future developments. The manager tries to anticipate trends and align the company with change or actively shape it through proactive as well as strategic leadership.

As a **Talent Scout & Talent Developer**, it is the responsibility of a manager to recognize and develop the potential of employees. The shortage of skilled workers and international competition make it more difficult to recruit suitable employees. Moreover, the continuous development of employees is becoming increasingly important in view of advancing technological progress.

As a **Culture Manager**, a manager establishes a corporate culture and a working environment in which the company's goals can be realized. The manager anchors methods and sets the framework to create a pleasant, productive, and inspiring work atmosphere.

The seven roles of modern managers have been derived based on a literature review in step 1. Figure 2 shows a cutout of the most significant correlations between the megatrends and the management roles. It is evident that all roles have a trend-oriented focus. Especially the megatrend digitalization plays a significant role.



megatrends roles	Digitalization	Globalization	Volatility	Sustainability	Demographic change
Data Analyst	large amounts of data available	global data streams in a networked world	identify trends faster based on data		
Change Manager	need for change in digital areas		change processes are on the daily agenda	ecological transformation leads to fundamental changes	changed understanding of values; involve employees
Digital Communicator	new and digital forms of communication are becoming established	employees and customers are globally dispersed			young and digital employees demand new forms of communication
Influencer & Digital Role Model	external presentation of managers is shifted to digital space	companies and their managers must show global presence		sustainability as a key topic in the public image of companies	young and digital employees who want to be inspired
Trend Seeker & Visionary	proactively shaping digital change		recognizing trends and developments and shapi	shapin	
	qualified employees in				

Figure 2: Correlation matrix of preliminary roles and megatrends

In the survey, the participating managers were asked to assess the relevance of each management role. The importance of the roles in the personal workday is widely confirmed, and the importance of the seven roles rises in the future. Figure 3 shows the comparison between the evaluated relevance of each role right now (black markers) and for the future (blue markers) (in 3 to 5 years). Thereby, the relevance of the management roles in the future personal work situation is consistently rated higher. The most significant discrepancy between the actual and the future situation is evident in the role of an Influencer & Digital Role Model. In total, the roles of Talent Scout & Talent Developer, Change Manager, and Trend Seeker & Visionary are currently the most important in the personal day-to-day work of the respondents and will continue to be so in the future. As a further result of the survey, competence profiles were derived for the respective roles. Figure 4 contains the five most essential competences for each role as rated by the survey participants, as described in chapter three.

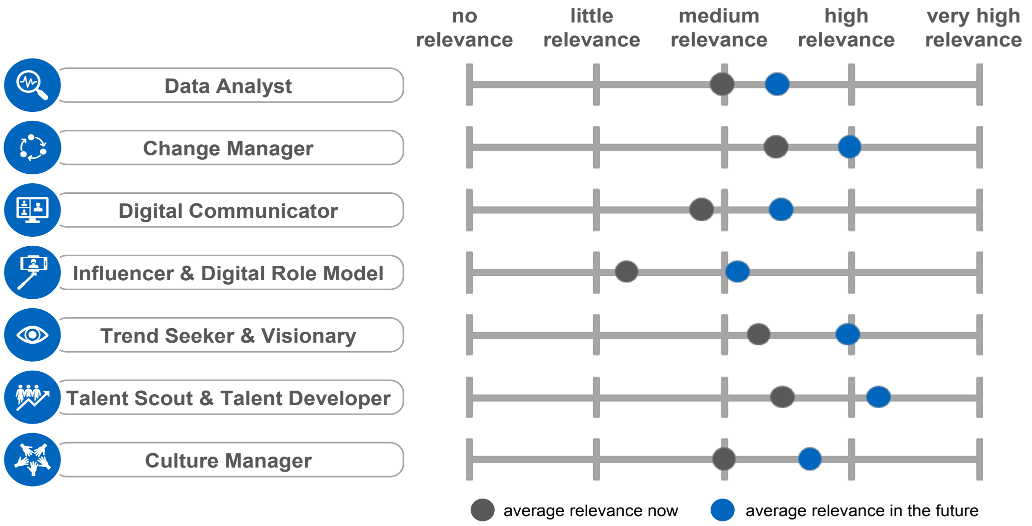


Figure 3: Assessment of the current and future relevance of the individual management roles

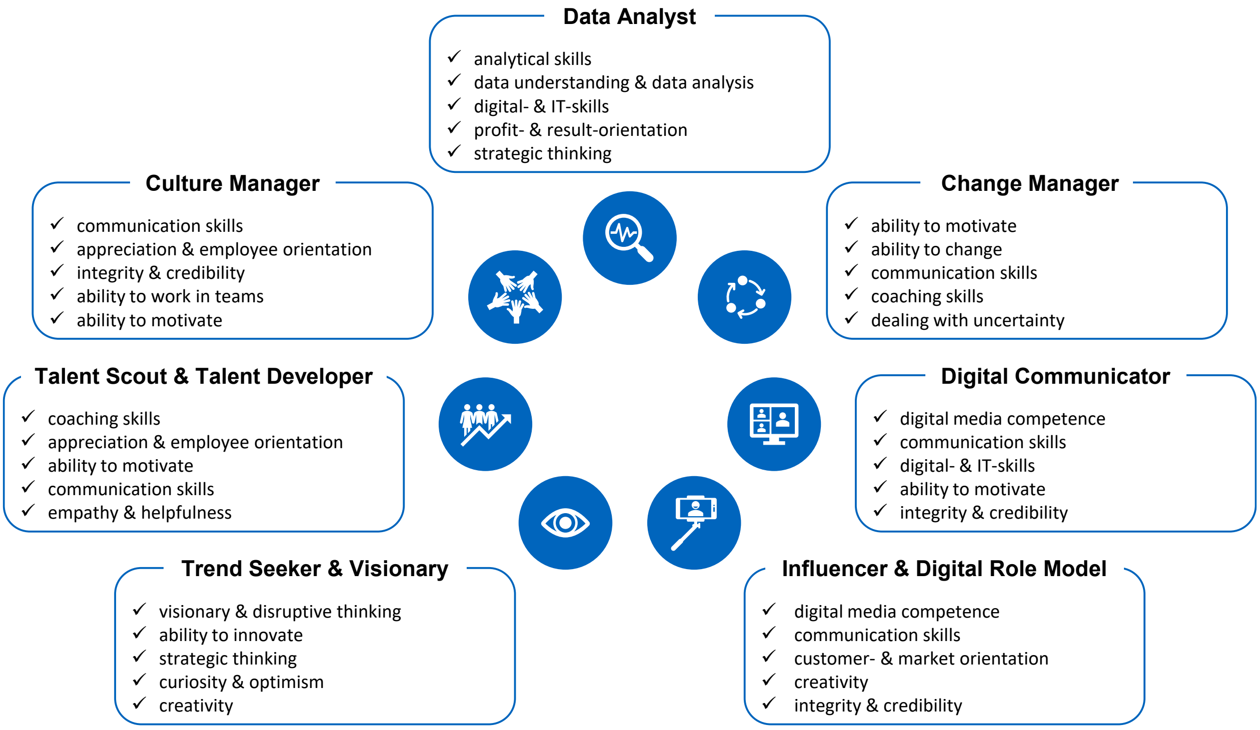


Figure 4: Seven roles of modern managers with the associated five most important competences

5. Interpretation and discussion

The developed competence-based role model provides a wide range of possible applications. The model reflects managers' central roles and competences in the context of current megatrends. Therefore, managers can use the competence-based role model to analyze their personal competence profile for their self-assessment. Based on the seven roles of modern managers, it is possible to assess to what extent and in what way the different roles are fulfilled and where the further potential for development exists. In addition, the competence-based role model can be used for the external assessment of managers. Possible applications arise when filling new management positions. For example, the model can serve as a starting point for developing a job profile. Although the 60 participants represent different company divisions, age groups, and roles, it should be noted that they do not constitute a holistically representative sample. Therefore, the study results of the questioned managers are only transferable to a limited extent. In addition, the competence profiles of the roles were developed based on the results of the empirical survey. They reflect the average opinion of the survey participants.

6. Conclusion

Nowadays, the day-to-day tasks of managers are highly complex and dynamic, as the environment for manufacturing companies is constantly changing in the context of digital transformation and current megatrends. Consequently, the required competences of managers and their roles are transforming. New competence-based role descriptions for managers are therefore necessary. In this paper, a competence-based role model of managers in manufacturing companies was presented, considering the influence of current trends. First, seven roles of modern managers were derived analytically and detailed based on role descriptions. In the course of an empirical survey of 60 managers from manufacturing companies, the roles were validated and analyzed in terms of their increasing relevance. Subsequently, in the context of the study, competence profiles were created for the respective roles with the help of a generated competence collection. Thus, the model describes and illustrates managers' central tasks and competences in role descriptions that arise from current megatrends. Future research is needed to transfer the findings into practice. To increase the benefits and possible applications, practical guidelines and trainings should be developed, enabling managers to develop their future competences in a more targeted manner.

Acknowledgments

We want to thank the Bavarian Research Foundation (BFS) for funding this work as part of the research project "Strain and competence-oriented employee scheduling" (BeKoMi, AZ-1336-18).

References

[1] Corejova, T., Chinoracky, R., 2021. Assessing the Potential for Digital Transformation. Sustainability 13 (19), 11040.

[2] El-Bassiouny, N., 2019. 20 Megatrends Shaping Management Thought in 2050, in: , 7th International OFEL Conference on Governance, Management and Entrepreneurship: Embracing Diversity in Organisations. April 5th - 6th, 2019, Dubrovnik, Croatia. Zagreb: Governance Research and Development Centre (CIRU), pp. 219–223.

[3] Mack, O., Khare, A., Krämer, A., Burgartz, T., 2016. Managing in a VUCA World. Springer, Heidelberg, New York, 259 pp.

[4] Jodoin, S., 2017. The Changing Role of Managers. https://www.linkedin.com/pulse/changing-role-managers-susan-jodoin. Accessed 17 March 2022.

[5] Koontz, H., O'Donnell, C., 1955. Principles of Management: An analysis of managerial functions. McGraw-Hill Book Company, New York.

[6] Rühli, E., 1985. Unternehmungsführung und Unternehmungspolitik 1. Haupt, Bern.

[7] Mintzberg, H., 1973. The nature of managerial work. Harper & Row, New York.

[8] Quinn, R.E., 1988. Beyond rational management: Mastering the paradoxes and competing demands of high performance. Jossey-Bass, San Francisco.

[9] Quinn, R.E., Rohrbaugh, J., 1983. A Spatial Model of Effectiveness Criteria: Towards a Competing Values Approach to Organizational Analysis. Management Science (3), 363–377.

[10]Lanwehr, R., 2013. Balancemanagement als Erfolgsfaktor von Organisation und Führung, in: Mai, D. (Ed.), Balance Management. Vom erfolgreichen Umgang mit gegensätzlichen Zielen. Springer, Wiesbaden, pp. 9–33.

[11]Denison, D., Hooijberg, R., Quinn, R., 1995. Paradox and Performance: Toward a Theory of Behavioral Complexity in Managerial Leadership. Organization Science (5), 524–540.

[12]Reinhart, G., 2017. Handbuch Industrie 4.0: Geschäftsmodelle, Prozesse, Technik. Hanser, München, 774 pp.

[13]Zeichhardt, R., 2018. E-Leader, CDOs & Digital Fools – eine Führungstypologie für den digitalen Wandel, in: Wassef, R. (Ed.), Disruption und Transformation Management. Digital Leadership - Digitales Mindset - Digitale Strategie. Gabler, Wiesbaden, pp. 3–21.

[14]Graf, N., Rascher, S., Schmutte, A.M., 2020. Teamlead – Führung 4.0: So führen Sie Teams synergetisch zu Höchstleistungen – Mit Tipps & Checklisten für die Praxis. Springer Gabler, Wiesbaden.

[15]Huynh, T.N., Hua, N.T. an, 2020. The relationship between task-oriented leadership style, psychological capital, job satisfaction and organizational commitment: evidence from Vietnamese small and medium-sized enterprises. JAMR 17 (4), 583–604.

[16]Willaert, P., van den Bergh, J., Willems, J., Deschoolmeester, D., 2007. The Process-Oriented Organisation: A Holistic View – Developing a Framework for Business Process Orientation Maturity, in: Alonso, G. (Ed.), Business process management. 5th international conference, BPM 2007, Brisbane, Australia, September 24-28, 2007; proceedings. Springer, Berlin, Heidelberg, pp. 1–15.

[17]Heyse, V., Erpenbeck, J., 2009. Kompetenztraining: Informations- und Trainingsprogramme. Schaffer-Poeschel, Stuttgart, 727 pp.

[18]IFIDZ, 2019. Führungskompetenzen im digitalen Zeitalter: Eine Analyse von 61 Studien und Umfragen aus den Jahren 2012-2018.

[19]Stifterverband, McKinsey, 2018. Future Skills: Welche Kompetenzen in Deutschland fehlen.

[20]Cloots, A., 2020. Digitale Kompetenzen: Welche es braucht und wie man sie erlernt, in: Wörwag, S., Cloots, A. (Eds.), Human digital work - eine Utopie? Erkenntnisse aus Forschung und Praxis zur digitalen Transformation der Arbeit. Springer Gabler, Wiesbaden, pp. 257–268.

Biography

Barbara Tropschuh (*1994) is currently a Ph.D. candidate at the Institute for Machine Tools and Industrial Management (*iwb*) at the Technical University of Munich (TUM). She received her master's degree in Mechanical Engineering from TUM and is now leading the research field of "human in production". Her research activities focus mainly on competence- and strain-oriented employee scheduling.

Benedict Beck (*1996) is currently a M.Sc. candidate at the Institute for Machine Tools and Industrial Management (*iwb*) at the Technical University of Munich (TUM). He received his bachelor's degree in Mechanical Engineering from TUM. His research activities focus mainly on competences in the context of digital transformation.

3rd Conference on Production Systems and Logistics

Technology Adoption Of Collaborative Robots For Welding in Small And Medium-sized Enterprises: A Case Study Analysis

Simon Schumacher[1], Roland Hall[1], Anna Waldman-Brown[2], Lindsay Sanneman[3]

[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstr. 12, 70569 Stuttgart, Germany*

[2]*Industrial Performance Center, Massachusetts Institute of Technology, 400 Main Street (Building E19-733), Cambridge, MA 02142, USA*

[3]*Department of Aeronautics and Astronautics, Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139, USA*

Abstract

Welding tasks in small and medium-sized enterprises (SMEs) are exemplary for high mix and low volume manufacturing. Today, 96 % of newly installed workplaces in SMEs are set up without a robot, although collaborative robot solutions for small lot sizes are emerging. In a study across the institutional ecosystem, technology providers, and technology adopters of the regional state Baden-Württemberg in Germany, SME technology adoption with respect to the use of collaborative robots was surveyed in form of expert interviews with feedback by direct users. The study helps SMEs to understand the necessary requirements and prerequisites for the design of collaborative work systems. As a result, the main barriers and potentials of the use of collaborative robots in welding in SMEs are presented, including ergonomic benefits to workers, the importance of skilled tradesmen in robot programming, and the lack of general robot knowledge across SMEs. Furthermore, the detailed analysis of two case studies gives insights into individual implementation processes at pioneering SMEs in this technological application field.

Keywords

Industrie 4.0; Human-Robot Interaction; Collaborative Robots; Technology Adoption; Small and Medium-sized Enterprises.

1. Introduction

For over ten years, Industrie 4.0 has held various promises for technological advances with regard to productivity, cost-efficiency, quality and flexibility improvements [1,2]. We define Industrie 4.0 as the existence of networked factory systems that use “intelligence” from sensors and algorithmic decision-making to collaborate with other machines and humans within the factory, rather than operating independently [3]. Larger manufacturing companies have implemented multiple use cases based on Industrie 4.0 technologies, while significant effects on productivity outcomes could not always be observed [4,5]. Although they represent a valuable fraction of industrial economies, small and medium-sized enterprises (SMEs) are especially lagging behind in terms of Industrie 4.0 technology adoption [6]. Large manufacturers are equipped with a multitude of industrial robots, and installations have grown steadily over the past years as robots have become more capable of both intelligence and networking across other factory systems [7]. SMEs have not yet reached the same level of automation due to limitations in economies of scale and investment budgets.

SMEs are characteristically high mix and low volume manufacturing, which provides viable opportunities for the use of Industrie 4.0 technologies for increased flexibility [8], such as smarter and easier-to-program

DOI: https://doi.org/10.15488/12176

robots that can support human-robot collaborations [3]. This study aims to explore Industrie 4.0 technology adoption with respect to the use of collaborative robots in welding processes at SMEs. The study is designed across the institutional ecosystem, technology providers, and technology adopters of the regional state Baden-Württemberg in Germany.

This paper presents the findings from the study and is structured as follows. Chapter 2 explains the relevant backgrounds for collaborative robot welding in SMEs, including this study's definition of SMEs, the basics of human-robot collaboration and the manufacturing ecosystem of collaborative robot welding. In chapter 3, the study design and research methodology are presented. Chapter 4 encompasses the overall results of the study with technology adoption barriers and benefits, as well as two case studies with detailed information on the technology adoption at two representative SMEs. The limitations of the study are discussed in chapter 5, followed by the conclusions in chapter 6.

2. Background: the use of collaborative robots in welding processes at SMEs

SMEs represent 99 % of all businesses in the EU as well as in Germany [9,10]. SMEs are usually characterized by the headcount of employees, sometimes in combination with financial limits. According to the definition of the European Commission, SMEs are defined by a staff headcount of up to 250 employees and either a turnover of €50 million or below or a balance sheet total of €43 million or below [9]. In addition to SMEs, the widely used but not clearly defined German term *Mittelstand* exists, which describes highly focused, very efficient and often family-owned enterprises of up to 500 employees [11,12]. For this study, the definition limit for an SME was set to up to 500 employees.

As one of many technological advances of Industrie 4.0, the use of collaborative robots in the form of human-robot collaboration (HRC) or human-robot interaction (HRI) has become a broad spectrum for research activities [13–16,3]. Human-robot collaboration has the potential to safely increase productivity of human labour and improve the ergonomics of manual tasks, by optimizing for the inclusion of a human participant in the decision-making loop as a member of a human-robot team [17]. A collaborative robot is a robot that is capable of collaborative operation, defined as an operation where purposely designed robots work in direct cooperation with a human within a defined workspace [18]. Thus, a collaborative operation is always defined by a combination of workspace and task specifics, resulting in four different interaction levels according to Behrens (2019): (1) shared workspace without shared task; (2) shared workspace and shared task without physical interaction; (3) Shared workspace and shared task that is ''handed-over'' from human to robot; and (4) shared workspace and shared task with physical interaction [19,20]. Users increasingly expect cobots to be easier to program than industrial or non-collaborative robots, often by the same shop-floor workers who share workspace with the cobots rather than dedicated robot programmers.

The range of applications for use of collaborative robots includes assembly operations, transportation of goods, material handling and commissioning, machine feeding, service robotics, and the automation of unergonomic tasks, e.g. in welding. Welding is a manufacturing process to join materials, e.g. metals, by using high heat to melt different parts together and allowing them to cool, causing fusion [21]. Welding can be carried out with different filler materials and energy sources. In combination with a welding nozzle on a collaborative robot arm, different gas welding types are available [21].

This paper is concerned with an analysis of the manufacturing ecosystem and value chain of collaborative robots for welding, which is displayed in Figure 1. The value chain consists of technology providers and technology adopters as well as the end customers. The central group of technology adopters in this study is only represented by the sub-group of SMEs of up to 500 employees. In general, all types of metalworking manufacturing companies qualify as technology adopters and there are various examples for applications in larger manufacturing enterprises. The supplier side is represented by multiple stages of technology providers. The collaborative robots are produced by the original robot manufacturers, and then either sold via regionally distributed resellers and integrators, or via a joint reseller/integrator organization. The end customers are not central in this study and have only been included as an external group with certain needs and requirements towards the technology adopting SMEs. The ecosystem of collaborative robot welding includes third-party institutions, associations and other funding bodies providing valuable assistance for SMEs and the other value chain entities. As new solutions such as easy programming becoming available, start-ups and other innovative service providers gain importance in the manufacturing ecosystem for collaborative robot welding.

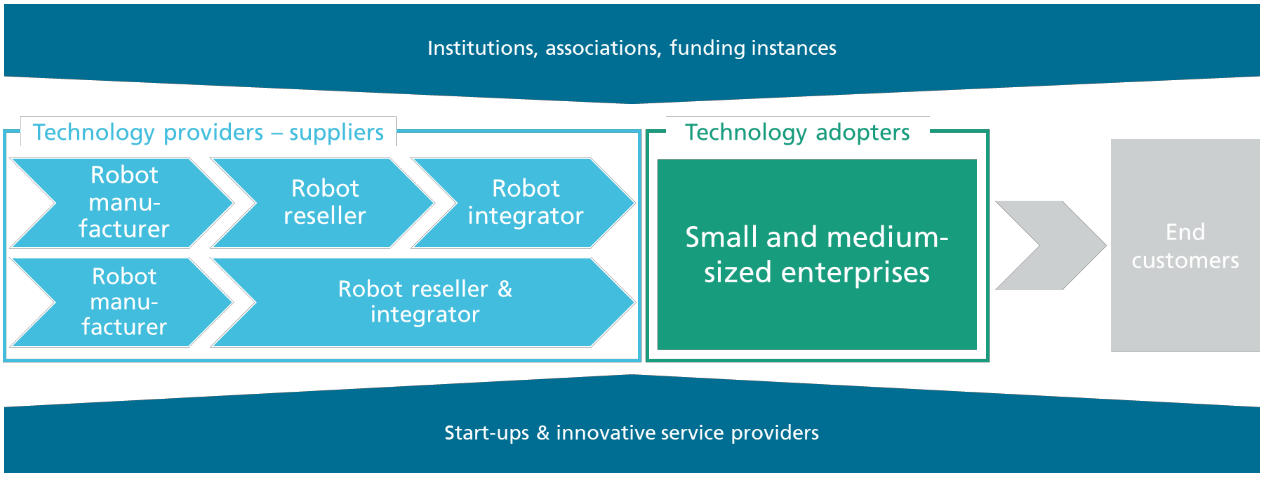


Figure 1: Manufacturing ecosystem for collaborative robot welding

3. Study design and research methodology

In the tradition of firm-level corporate interviews and factory tours across sociology and human factors research [22,23], this study consists of a field study with interviews of participants in the collaborative robot welding ecosystem of the German regional state Baden-Württemberg, from November 2021 – January 2022. The aim of the study is to analyse the Industrie 4.0 technology adoption of SMEs as well as the institutional ecosystem. According to the manufacturing ecosystem presented in Figure 1, the study covers four levels: (1) institutional ecosystem; (2) technology providers; (3) technology adopters; (4) workplace and workers. As mentioned beforehand, the study is focused on SMEs as technology adopters. The workplace level is included in the study in order to not only analyse the managerial decision making by SMEs, but also the potential shop floor changes through the introduction of new technologies and their implications for new skills and competence profiles.

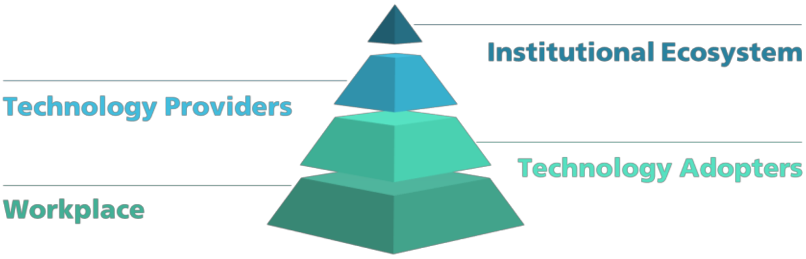


Figure 2: Study levels

The study design as well as the execution of interviews were affected by the global coronavirus pandemic. Interview questionnaires were designed to be used for both in-person and remote interviews. Strict General Data Protection Regulation (GDPR) measures were applied in order to guarantee the anonymous and only study-related use of company data during and after the interviews. E.g. the interviewees were always free to quit the interview or not to answer certain questions, and the raw data was held on secured servers reserved for this study only.

The selection of participants for the study was made upon certain criteria to reach representative participants from each of the four study levels. Public and private ecosystem institutions were required to provide active support for SMEs to assess and acquire new Industrie 4.0 technologies. Criteria for technology providers were the provision of collaborative robots in an original manufacturer, reseller, or integrator role with multiple products sold to SMEs. Criteria for SMEs to be assessed as technology adopters were the workforce limit of up to 500 employees, the implementation of a cobot solution and the access to workplace insights. Because of the scarcity in SMEs with cobots for welding automation, the selection of participants was not made randomly, but through mainly publicly available information on pioneering SMEs with early installations. The analysis of the interviews from this field study is mainly based on qualitative data to explore the Industrie 4.0 technology adoption of the sample group. The analysis process is based on video and audio recordings, anonymised transcription of the recorded interviews, and attribution of codes for statements. A common transcription software was used for the transcription and coding activities. The clustering of codes allowed for detailed analysis of the interviews. The qualitative interview data serves to summarize technology adoption barriers and benefits for the sample group of interviewed participants.

4. Study results: Industrie 4.0 technology adoption by SMEs in Baden-Württemberg

During the field study from November 2021 – January 2022, 16 interviews were carried out with companies and institutions from Baden-Württemberg. The field of participants consists of 7 public ecosystem institutions, 1 private ecosystem start-up, 4 technology providers (who have collectively provided dozens of collaborative robots to SMEs), and 4 technology-adopting SMEs (representing a total of 862 workers who have been exposed to cobots, although only a handful of workers regularly interact with the robots). Each group was interviewed with an adapted semi-standardized questionnaire, which asked how companies decided to buy a robot, how they integrated the robot, how they chose and educated people to operate the robot, whether operation has differed from expectations, and how the robot itself might be improved. The interviews were recorded, transcribed and analyzed. All technology providers and adopters voluntarily provided access to their factories and workplaces. In this chapter, the results will be presented in an overall summary of technology adoption barriers and benefits deduced from all 16 interviews. Sections 4.1 and 4.2 provide detailed insights into the technology adoption at SMEs through two case studies from representative companies, in which many of the barriers and benefits listed below can be identified.

In general, the reported barriers for cobot welding adoption are relatively low– in stark contrast to the difficulties experienced by SMEs adopting industrial robots [6,8]. All participants of the ecosystem and especially the SMEs as users of the technology were satisfied or even surprised by the successful implementation process. Open-mindedness was mentioned the most as a crucial success factor or barrier, when missing. Table 1 lists further technology adoption barriers, especially with regard to the current limitations of the relatively new technology.

Table 1: Summary of technology adoption barriers from the interviews

Technology adoption barriers for welding cobots
1. Mindset: Management and workers cannot be closed-minded about new technology
2. Lack of cobot versatility: Humans can better adjust to unexpected events and perform complex/critical tasks
3. Lot sizes: Cobot automation of jobs is rarely beneficial for very small (<5) and large lot sizes (>100)
4. Physical interference: Cobot can only weld in certain positions without bumping into itself or other limitations
5. Welding jigs: Cobots require new jigs to hold parts in place during robotic welding, while humans can use hands
6. Monotonous leftover jobs: After partial automation, cobots may create boring and monotonous leftover tasks

The interviews shed light on various possible benefits of the Industry 4.0 technology adoption for SMEs. The introduction of cobot welding has several positive effects on the work system, e.g. improvements of human factors and ergonomics for workers [14,16]. The work organization allows the welding experts to be autonomous in terms of their decision of which work piece to weld manually or with cobot assistance. The ease of use even causes excitement for the job, and was mentioned by two firms as a selling point for attracting welding apprentices to adopting workplaces. Most of the process preconditions are fulfilled by the cobot, while output in terms of high and steady welding quality as well as reduced rework can be optimized. From a managerial standpoint, the relatively low investment cost seems to be over-compensated by savings for health insurance costs, reduced wages for contract welders, waste reduction, and flexibility gains in shift planning. See Table 2 for the list of technology adoption benefits from the interviews.

Table 2: Summary of technology adoption benefits from the interviews

Technology adoption benefits for welding cobots
1. Human factors: Less toxic fumes breathed in by welders
2. Human factors: Reduced stress for eyes, particularly for older workers
3. Human factors: Reduced physical strain with regard to uncomfortable posture, particularly for older workers
4. Task: Welding is a very suitable process for cobot automation due to tooling and relatively low feeding speed
5. Quality: Cobots provide consistently high-quality welding results over a whole shift or longer
6. Rework: Cobot welding can reduce the overhead on rework processing due to the homogeneity of welds
7. Ease of use: Cobot programming is perceived as very easy to use by the welding experts
8. Investment costs: Relatively low investment costs (ca. 50.000 € – 100.000 €) versus industrial robots
9. Operating costs: Welding cobots can be cheaper than workers for repetitive, non-variable tasks
10. Worker shortages: Automation of a fraction of welding jobs can help to cope with welding expert shortages
11. Flexibility: Cobot use can be scaled up or down instead of hiring contractor workers
12. Added shifts: Cobot can work through night shifts, holidays, and when humans are unwilling to work
13. Worker autonomy: Welders can choose on their own which jobs to automate and which to do manually
14. High-tech signalling: Adopting firms can use cobots to attract new workers and advertise to end customers

Considering the task profile of human, collaborative or automated welding, different variables have to be taken into account when planning for the right individual workplace setup. In combination with other variables such as the geometry and length of welding parts and seams, the interviewees named the lot size of planned jobs as an important variable for the use of collaborative robots for welding. Even when jobs are technically feasible for welding with a collaborative robot, the process of programming the cobot for a small batch size takes too long to justify automation. In SMEs, this planning task is usually transferred to the welding expert with cobot programming skills, who decides which jobs to automate. From the interviews, these experts seemed to appreciate the additional workforce provided by the cobot. With growing, but rather low reported lot sizes, the programming efforts scale with the lot size. Since SMEs typically have high variety in their products and jobs, the lot size should be considered as a key variable in planning of cobot welding capacities. For higher lot sizes, cobot welding faces competition with existing industrial robot solutions, which tend to outperform them at scale even though the initial programming takes longer. However, other variables apply as well and need to be considered holistically.

4.1 Case study 1: Medium-sized enterprise with movable welding cobot

The first case study is based on the interview with the head of production of an SME from Baden-Württemberg with 220 employees and annual revenues of about €23 million. The SME produces make-to-order goods with smaller lot sizes. The SME has global customers, four international factory locations, and delivers in the premium quality segment for special applications. The SME covers the full value-creation chain from electromechanical engineering and constructing, manufacturing, assembling, mounting, and servicing for the produced goods. With regard to the institutional ecosystem, the SME is member of the Chamber of Commerce and Industry (*Industrie- und Handelskammer, IHK*) and the German Welding Society (*Deutscher Verband für Schweißen und verwandte Verfahren e.V., DVS*). The SME regularly hosts three apprentices per year in its own facilities and is actively involved in the apprenticeship curriculum in collaboration with IHK and DVS. The SME also collaborates with research institutions for specific engineering-related projects and opens up its factory as a best practice for visits by other German manufacturing companies.

As one of the first users of this technology, the SME set up a workplace consisting of a *Universal Robot* collaborative robot arm equipped with a welding tool and nozzle by an integrator. The workplace serves as an addition to other manual welding stations. In the initial setup from 2018, the cobot was attached to the welding table and able to be moved to various positions– which turned out to be too rigid (Barriers #2 and 5). In order to build two separate workplaces next to each other, a five meter long linear axis was installed above the table with a hanging cobot solution in 2019. This new and current setup allows for the next job to be prepared by a human co-worker while another welding job is running. Usually, the preparation covers the removal of finished goods, the optional change of welding jigs, and the positioning of new welding pieces. Figure 3 and 4 show manual welding and the current cobot welding workplaces.

The cobot welding solution was first identified as an innovation at a metalworking trade fair in 2017. With a suitable job of 100 identical parts to be delivered in 2018 (Barrier #3), both middle management and the owner were convinced that the relatively low investment costs would quickly pay off by freeing up more time for valuable human welders (Benefit #9) and improving workplace satisfaction by reducing welding fumes (Benefit #1). Except for welding time calculations, no other investment or cost-related calculations were made: the cobot's welding speed was comparable to human welders (Benefit #4). Due to the steel types welded at this SME, it was crucial that the cobot welding solution acquired was capable of switching between metal inert/active gas (MIG/MAG) and tungsten inert gas (TIG) welding, which was satisfied by a market ready and integrated solution. The trust as well as the open-mindedness by the owner/CEO is typical for an SME and was important for the fast acquisition process (Barrier #1). The welding manager estimated that 95 % of the firm's welding jobs could theoretically be done by the cobot, yet the cobot is only responsible for about 3 % of jobs due to the burden of programming time for smaller batches.

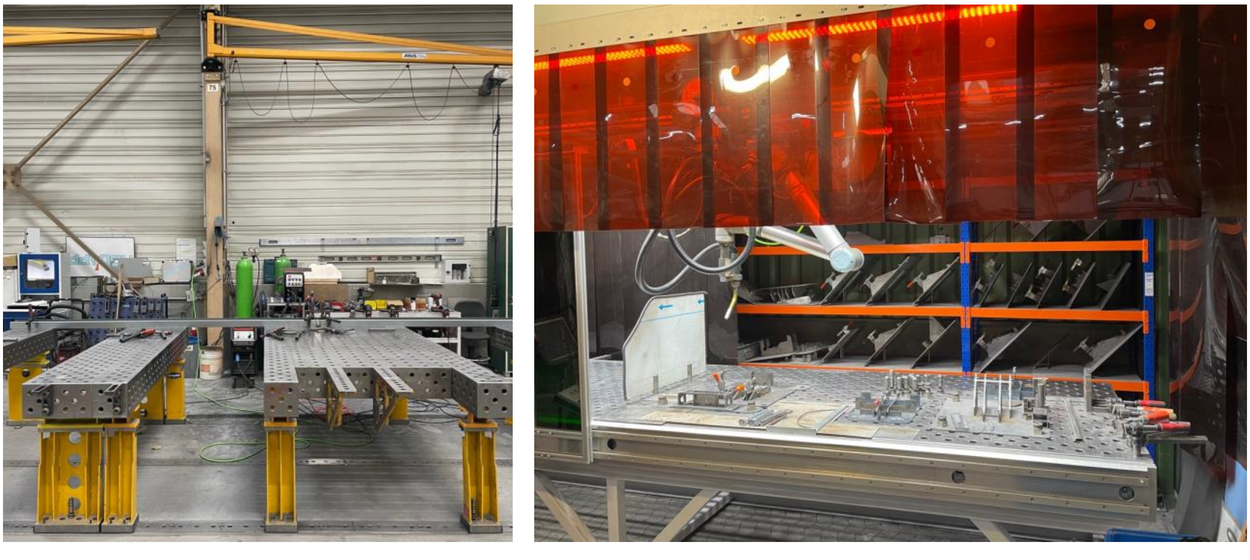


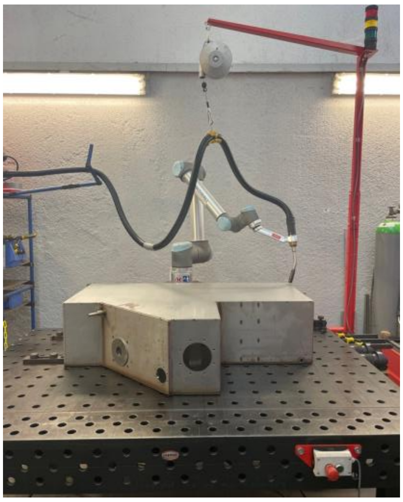
Figure 3 and 4: Workplaces for manual welding (left) and cobot welding (right) at the SME from case study 1

The SME identified further use cases for the application of a second cobot workplace, which is already planned as a next acquisition. The identified case would use another Universal Robots cobot to handle materials in an integrated IT system with an existing laser cutting machine, and could be used to add a night shift in addition to the current work organization (Benefit #12). The company is also trying to advance with the current welding cobot solution to weld thinner metal sheets as well as to add sensors to the welding tool for higher adaptability and quality (Benefit #5).

4.2 Case study 2: Medium-sized enterprise with stationary welding cobot

The second case study is based on an interview with three representatives of an SME from Baden-Württemberg with 160 employees and annual revenues of about €26 million. The SME produces individually engineered machinery for a special branch application. The SME exports 80 % of its goods to global customers and offers a product range from small special machinery to small businesses, as well as premium and huge-dimension special machinery to industrial corporate firms. The SME covers the full value-creation chain from electromechanical engineering and constructing, manufacturing, assembling, mounting, and servicing for the produced goods. Over 70 current employees, i.e. nearly 50 % of the workforce, completed work-study apprenticeships in this company, and the SME hosts multiple apprentices per year in its own facilities. With regard to the institutional ecosystem, the SME as an early pioneer user collaborates with technology providers and research institutions in order to identify new improvement measures for the cobot welding workplace. Since 2007, the SME has built up and established its own lean production system, which is not typical for an SME, including a Kaizen system for continuous improvement processes, and various Lean principles and methods such as a milk run based on a Kanban system.

As a pioneering SME, this company acquired a first version of an integrated cobot solution provider in 2017 before its official market entry. The cobot is placed in an upright position on a typical welding table and can be flexibly moved (Barrier #4). As required by the SME from case study 1, the firm needed a solution to switch between the welding types MIG/MAG and TIG. With improvements during the last years, the analysed setup can be used for welding stainless steel plates of 1 mm or thicker. The cobot welding results were reported to be of high overall quality, leading to drastic reductions in rework after welding (Benefits #5 and 6). Figure 5 and 6 show the cobot welding setup with an exemplary, relatively complex work piece.



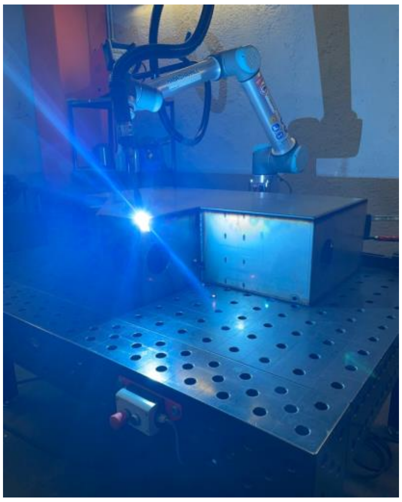


Figure 5 and 6: Workplace for collaborative robot welding at the SME from case study 2

The selection of programmers for the cobot was dependent on existing proficiency as a welding expert as well as open-mindedness (Barrier #1). At this SME, three welding experts were qualified for this role. One of them is under 22 years old, while the oldest is over 60 years old and appreciates the reduced eye strain from programming the robot instead of welding himself (Benefit #2). The training provided by the integrator took less than one day, followed up by 1-2 weeks of workers figuring out the programming on their own.

The direct users were surprised by the ease of use and improved ergonomics, and found joy in this new task (Benefits #3 and 7). The introduction of this technology was accompanied by the works council and did not lead to changes in wages paid, but to up-skilling and increased involvement of welders (Benefit #13). The SME built up expertise in their own construction of jigs and racks for welding automation (Barrier #5). In an exemplary joint effort by welding experts and the Lean production team, a high-volume job of 300-400 pieces per year was set up for cobot welding, using a rack for five pieces per automated welding program.

With help from shop floor welders, production managers have identified further work pieces to be welded by means of the cobot. Furthermore, the SME engages with research institutions to integrate sensor-based features for automated weld seam following, which could enable further applications and the need for standardisation of fabricated parts (Barrier #2). The management did not calculate the overall economic effects, but is convinced that the solution has led to several improvements for both workforce and company. The ease of use of cobot programming as well as the open-mindedness by the workers were stressed multiple times by the interviewees.

5. Limitations of the study

Although this study presents new insights, limitations are inherent in any study and should be transparently addressed. Due to the need to interview SMEs with the cobot technology in use, participants were not randomly selected. Another sample group could have resulted in different and further findings. The sample group of 16 participants including only 4 SMEs does not allow for broadly-applicable conclusions. As laid out in the research methodology, it was not planned to derive quantitative data from the interviews, but to perform a qualitative analysis on the ecosystem. The results from chapter 4 are promising for this approach. The authors look forward to continuing with interviews on this topic in order to gather further insights from the manufacturing ecosystem in an international context, especially focusing on the contribution of a publicly available qualitative and quantitative data set (see outlook in chapter 6). The interviews were carried out by four researchers from two research institutions from different continents and from three different professions. Despite the diverse backgrounds, researcher bias can be assumed. A predefined, semi-standardized questionnaire with a fixed set of questions was used to standardize and objectify the interviews. As another limitation, the interviews as well as the analytical work were carried out in German and English, which led to minor difficulties in the general understanding of translated technical terminology as well as colloquial expressions. The overall analysis results should not be affected by this.

6. Conclusions

This paper shows the results of a study with different participants of the collaborative robot welding ecosystem in Baden-Württemberg, Germany with a focus on SME technology adoption. In 16 interviews, several technology adoption benefits were identified, including worker safety, worker autonomy, quality and performance improvements, and organizational flexibility. Although technology adoption barriers exist, the reported limitations are mainly based on technological limitations of cobots, which can be assumed to be advancing with future developments, and the often-stressed managerial mindset of open-mindedness. The interview results highlight the feasibility of cobot welding solutions even for SMEs as a technology adoption group. Two case studies give detailed insights into the barriers and benefits experienced by early adopting SMEs, who proved to be successful pioneers of cobot welding in the analysed region. With first hand impressions of the workplace level, the perception of this new technology by SME workers can be described as very positive, with welding professionals easily becoming cobot programmers based on the latest market-ready solutions.

The paper contributes to both academia and industrial practice. In contrast to other studies in which Industrie 4.0 technologies are too expensive and/or underdeveloped for use in manufacturing firms and especially SMEs, this study presents a solid picture of contemporary Industrie 4.0 technology adoption across a small subset of early-adopter SMEs [8]. The findings are based on interviews across the manufacturing ecosystem in a single regional state. Further industrial policy findings will be published in an according outlet. For industrial practice, the interview results and case studies serve as blueprints for other SMEs thinking of the adoption of similar cobot solutions. The interviewed companies were assisted with further assistance for own reflection with regard to improvement potentials. This research is part of an international comparative study between the industry-heavy regional states of Ohio, United States of America and Baden-Württemberg, Germany, which aims to better understand the workforce implications of robot adoption, and to support both SMEs and their workers in technological advancement. Future work will entail further interviews with robot-adopting SMEs and institutions, as well as additional analysis to build up a thorough data set for applied research purposes.

Acknowledgements

This paper is part of the Future Work Lab pilot project, which is funded by the German Federal Ministry of Education and Research (BMBF) within the program "Innovations for Tomorrow's Production, Services, and Work" - »Future of Work« and controlled and implemented by the Project Management Agency Karlsruhe (PTKA). We would also like to thank the MIT Center for International Studies, the Interactive Robotics Lab, and the MIT Industrial Performance Center for support.

References

[1] Kluth, A., Schiffer, M., Fries, C., König, J., 2020. Influencing factors of the digital transformation on the supply chain complexity dimensions. Journal of Production Systems and Logistics 1 (1).

[2] Neugebauer, R. (Ed.), 2019. Digital Transformation. Springer Vieweg, Berlin.

[3] Schumacher, S., Pokorni, B., Hall, R., Bildstein, A., Hämmerle, M., 2021. Development of a Practical Orientation Guide with Industrie 4.0 Use Cases for Industrial Manufacturers. Proceedings of the Conference on Production Systems and Logistics CPSL 2021, 319–328.

[4] Tabrizi, B., Lam, E., Girard, K., Irvin, V., 2019. Digital Transformation Is Not About Technology. https://hbr.org/2019/03/digital-transformation-is-not-about-technology. Accessed 27 August 2020.

[5] ZoBell, S., 2018. Why Digital Transformations Fail: Closing The $900 Billion Hole In Enterprise Strategy. https://www.forbes.com/sites/forbestechcouncil/2018/03/13/why-digital-transformations-fail-closing-the-900-billion-hole-in-enterprise-strategy/#4f74e9207b8b. Accessed 27 August 2020.

[6] Sanneman, L., Fourie, C., Shah, J.A., 2021. The State of Industrial Robotics: Emerging Technologies, Challenges, and Key Research Directions. FNT in Robotics 8 (3), 225–306.

[7] International Federation of Robotics, 2021. World Robotics 2021 Industrial Robots. VDMA, Frankfurt/Main.

[8] Waldman-Brown, A., 2020. Redeployment or robocalypse? Workers and automation in Ohio manufacturing SMEs. Cambridge Journal of Regions, Economy and Society 13 (1), 99–115.

[9] European Commission, 2021. SME definition. https://ec.europa.eu/growth/smes/sme-definition_en. Accessed 5 January 2022.

[10] German Federal Ministry of Economics and Technology, 2016. German Mittelstand: Engine of the German economy: Facts and figures about small and medium-sized German firms. BMWi, Berlin.

[11] Audretsch, D.B., Lehmann, E.E., Schenkenhofer, J., 2018. Internationalization strategies of hidden champions: lessons from Germany. Multinational Business Review 26 (1), 2–24.

[12] Pahnke, A., Welter, F., Audretsch, D.B., 2021. Im Auge des Betrachters? Warum wir zwischen KMU und Mittelstand unterscheiden müssen. Institut für Mittelstandsforschung Bonn, Bonn.

[13] Bdiwi, M., Pfeifer, M., Sterzing, A., 2017. A new strategy for ensuring human safety during various levels of interaction with industrial robots. CIRP Annals 66 (1), 453–456.

[14] Hollerer, S., Fischer, C., Brenner, B., Papa, M., Schlund, S., Kastner, W., Fabini, J., Zseby, T., 2021. Cobot attack: a security assessment exemplified by a specific collaborative robot. Procedia Manufacturing 54, 191–196.

[15]Malik, A.A., Brem, A., 2021. Digital twins for collaborative robots: A case study in human-robot interaction. Robotics and Computer-Integrated Manufacturing 68.

[16]Ostrowski, A.K., Pokorni, B., Schumacher, S., 2020. Participatory Design for Digital Transformation of Manufacturing Enterprises. MIT Work of the Future Working Paper Series. https://workofthefuture.mit.edu/research-post/participatory-design-for-digital-transformation-of-manufacturing-enterprises/. Accessed 5 January 2022.

[17]Gombolay, M., Bair, A., Huang, C., Shah, J., 2017. Computational design of mixed-initiative human–robot teaming that considers human factors: situational awareness, workload, and workflow preferences. The International Journal of Robotics Research 36 (5-7), 597–617.

[18]International Organization for Standardization, 2016. Technical specification ISO/TS 15066: Robots and robotic devices - collaborative robots. ISO, Vernier, Geneva, Switzerland.

[19]Behrens, R., 2019. Biomechanische Grenzwerte für die sichere Mensch-Roboter-Kollaboration. Springer, Wiesbaden.

[20]Koch, T., Beck, J., Mayer, S., 2021. Systematic approach to consider HRC in early design phase. Procedia CIRP 100, 774–779.

[21]Groover, M.P., 2019. Fundamentals of modern manufacturing: Materials, processes, and systems. John Wiley & Sons Inc, Hoboken, NJ.

[22] Berger, S., 2006. How We Compete: What Companies Around the World Are Doing to Make It in the Global Economy. Doubleday, New York, United States of America.

[22] Schoenberger, E., 1991. The corporate interview as a research method in economic geography. The Professional Geographer, 43 (2), 180-189.

Biographies

Simon Schumacher (*1990) is a researcher at the research group Implementation Methods for Digital Production at the competence center for Digital Tools in Manufacturing at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart, Germany. At Fraunhofer IPA, he is the project lead of the Future Work Lab and the strategic coordinator of the research field technologies for human-centered production. As a PhD student, he is enrolled at the University of Stuttgart. As a visiting student, he has been enrolled at the Massachusetts Institute of Technology in 2020 in a collaboration with the MIT Work of the Future Task Force. Another research stay at the MIT Industrial Performance Center is planned for 2022.

Roland Hall (*1993) is a researcher at the research group Implementation Methods for Digital Production at the competence center for Digital Tools in Manufacturing at the Fraunhofer-Institute for Manufacturing Engineering and Automation IPA in Stuttgart, Germany.

Anna Waldman-Brown (*1989) is a PhD candidate in political economy at MIT's Department of Urban Studies and Planning, where she researches automation, inequality, and SMEs with the MIT Work of the Future Task Force and the Industrial Performance Center. She is a graduate of the MIT Technology and Policy Program and a Fulbright fellow.

Lindsay Sanneman (*1992) is a PhD candidate in the Department of Aeronautics and Astronautics at MIT and a member of the Interactive Robotics Group. Her interests primarily lie in large-scale optimization for planning and scheduling. She received her bachelor's degree from MIT in June of 2014 and completed her master's work on multi-level optimization techniques for the supply chain in June of 2018. Her current research addresses collaborative scheduling and optimization through automates preference elicitation and explanation in large human teams.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2022

3rd Conference on Production Systems and Logistics

Towards Smaller Value Creation Cycles: Key Factors And Their Interdependencies For Local Manufacturing

Pascal Krenz[1], Dominik Saubke[1], Lisa Stoltenberg[1], Julia Markert[1], Tobias Redlich[1]
[1] *Helmut-Schmidt-University, Hamburg, Germany*

Abstract

The unpredictable occurrence of a global pandemic and trade conflicts have currently shown us the fragility of global, industrial value chains. In contrast to this, local value creation structures have numerous potentials to meet present ecological, economic and social challenges (e.g. increasing the resilience of the manufacturing sector, reducing CO_2 emissions through smaller loops of value creation, empowering regional stakeholders). This paper presents a study on local manufacturing designed to achieve a better understanding of the internal systematics of value creation in a local context using a sensitivity analysis. By modelling the causal effects, the direct and indirect influences of internal and environmental factors of local production as well as their independencies can be shown. This in turn will enable scenario analyses that show possible developments for local production systems arising due to changing social, political and technological factors. In the future these options may aid in decision-making processes aiming at a sustainable circular economy.

Keywords

Local Manufacturing; Distributed Manufacturing; Re-Distributed Manufacturing; Urban Manufacturing; Value Chain; Sustainability, Sensitivity Analysis

1. Introduction

Recently, local value creation has been discussed as an instrument to reduce the increasing risks of complex global value chains (e.g. resource scarcity, trade barriers) and to expand the possibilities of sustainable production (smaller value cycles, empowerment of regional actors) [1]. Current societal trends and challenges (individualisation, sustainable consumption) and technological innovations (flexibilization of production systems, new communication technologies, smart systems) could potentially promote this development [2–4,1] . Thereby local value creation can have various forms: These may include regional value creation clusters (e.g. Hamburg Aviation or Life Science Nord in the Hamburg region), decentralised production sites of globally operating enterprises, value creation through local crafts and the participation of citizens through fab labs or makerspaces.

Reducing the size of the value creation cycles has the potential to improve the sustainability of product manufacturing [5]. If a value creation system is aligned with the local context in terms of product manufacturing, producers and demand, this forms a particularly good basis for small value creation cycles; local context meaning local production of goods, utilisation of local resources, addressing of local demands [6]. In this paper, we examine local production as a phenomenon that gravitates towards these three characteristics and can thus be distinguished from global value creation.

DOI: https://doi.org/10.15488/12128 publish-Ing.

2. Research Question and Motivation

There are different and diverse approaches looking at the phenomenon of local value creation: e.g. Distributed Manufacturing, Re-Distributed Manufacturing, Local Production, Urban Production. Thereby the focus of consideration varies [6]. This variety of perspectives offers a broad picture of the different forms and characteristics of local manufacturing to the reader. But at the same time, the reader does not achieve a clear understanding of the object of study as the different foci (technological, economic, social perspective) distort the actual relevance of the key factors of local manufacturing. As a result, the attributes of local manufacturing bear the risk of being under- or overvalued depending on the viewing perspective.

This paper presents a study that aims to contribute to a better understanding for the systematics of local manufacturing. A holistic perspective is taken on the object of study, in which political, social, economic and technological factors and their interactions are considered. Therefore, the research question is: Which internal and external factors of a local value creation system have a major impact on the implementation of local production, in the sense of:

(A) The local production of goods (on site)

(B) The use or inclusion of local resources in the production processes (equipment, actors, materials)

(C) And the fulfilment of local demands

3. Method

The study is based on a systematic analysis that is guided by the methods according to Vester for the assessment of complex systems [7–9]. Vester's sensitivity analysis [8–10] facilitates the determination of interdependencies between system dimensions in complex systems [8]. That way, options for a targeted development of the system can be described [8]. The approach not only considers direct connections between system factors but also indirect causal effects as well as feedback effects and self-reinforcing (or self-weakening) loops. The goal is to identify **representative patterns for the functioning of the value creation system** in order to finally show possibilities for development of the system.

Step 1 - System description: The system description serves to capture and also delimit the object of study [10]. The object of this study corresponds to systems of local value creation, which are characterized by the following attributes that can be viewed as dimensions [6]: **(A) Local production of goods**, **(B) Utilization of local resources (stakeholders, materials)**, **(C) Addressing of local demands**.

A value creation system aims to provide a material or immaterial service in a systematic and structured manner [11]. Value creation systems can be categorized as socio-technical systems [12]. They behave partly deterministic and partly probabilistic so that the predictability of their behaviour is limited [13][14].

Step 2 - Identification of influencing factors: To identify the influencing factors [10] about 90 texts were scanned that deal with different concepts of local manufacturing (e.g. Urban Manufacturing, Distributed Manufacturing, Re-Distributed Manufacturing) and thus take different perspectives regarding the object of study [6]. For the deeper analysis of the content those texts were chosen from which characteristic factors influencing local manufacturing could be derived. The choice of texts followed the principle of theoretical saturation [15]. The relevant influencing factors on local manufacturing were identified from the texts through a process of itemization, abstraction and condensation (further illustration of the approach, see [6]).

Step 3 - Modelling of the causal network: The identified factors are put together in a causal network to capture the systemic interaction between them. This is based on the idea that a value creation system – while it is not deterministic – still has an inner order that can be uncovered to better understand the underlying systematics [14]. In the causal network the type of the different influencing factors, their causal effects among each other and the development over the course of time is assessed [10].

Based on the literature review, the identified factors and their described interdependencies were transferred into a model that visualized the causal network using the software iModeler. The model was consolidated in moderated, interdisciplinary workshops (based on [10]) in order to achieve a realistic representation of the causal network. Through this process, the influences between factors that were often described as **indirect influences** in literature were **successively reduced to their direct causal effects**. Indirect effects between system factors were represented by causal chains based on the developed direct influences.

The differentiation of direct and indirect factors in the model was important in order to determine the strength of the causal effects. This was done by comparatively assessing the causal effects of all input factors of one particular system factor. Therefore, the partial effect of an input factor was categorized into levels of impact (low 5%, moderate 10%, relevant 20%, significant 35%, essential 50%). The sum of the influences of all input factors is limited to a maximum of 100%. Throughout the assessment of the factors' impacts, the sum was usually kept below the maximum of 100% to account for all influences not depicted in the model. Furthermore, the causal effects between system factors over the course of time were considered (short-term: 1 to 5 years, mid-term: 5 to 10 years, long-term: > 10 years). When several factors are connected in causal chains, the influence of indirect factors is calculated by multiplying the percentages of the influences along the chain. The influence of these indirect factors decreases depending on the distance. As a result, the strength of influences can be modelled more realistically by differentiating between direct and indirect factors.

Step 4 - Evaluation of the causal effects in the model: Based on the developed model, the influences between system factors were analysed to answer the research question listed in Chapter 2. The aim is to identify the most important internal and environmental factors influencing local manufacturing, which can be described by three dimensions: **(A) Local production of goods**, **(B) Utilisation of local resources**, **(C) Addressing of local demands**. These dimensions are represented in the model by five **main attributes** of local manufacturing: **Production at the place of need**, **Use of local (raw) materials**, **Implementation of production by local stakeholders**, **Production of individualised / locally adapted products**, **On demand production** (refer to Table 1). The relevance of the impact of the system factors on the five main attributes of local manufacturing was assessed by evaluating the strength of influence of the direct and indirect input factors. The result is a **comparative assessment** of those factors influencing the main attributes of local manufacturing, which will be presented in the following.

Table 1: Dimensions of local production and their depiction in the model

Dimension of local manufacturing	**(A) Local production of goods**	**(B) Utilisation of local resources**	**(C) Addressing of local demands**
Main attributes of local manufacturing	○ Production at the place of need	○ Use of local (raw) materials ○ Implementation of production by local stakeholders	○ On demand production ○ Production of individualised / locally adapted products

4. Findings

The model includes a total of 160 factors. The characteristics of local manufacturing are represented by five main attributes in the model (refer to Table 1). The model also includes target factors of a local production (i.e. prosperity in the region, sustainable production), current trends (i.e. digitalisation, dynamization of the markets/the environment of value creation, individualisation, urbanisation) as well as other relevant factors (of technological, political, economic and social influence). Figure 1 shows an exemplary section of the model depicting the main key attribute ***production at the point of need*** with selected interactions.

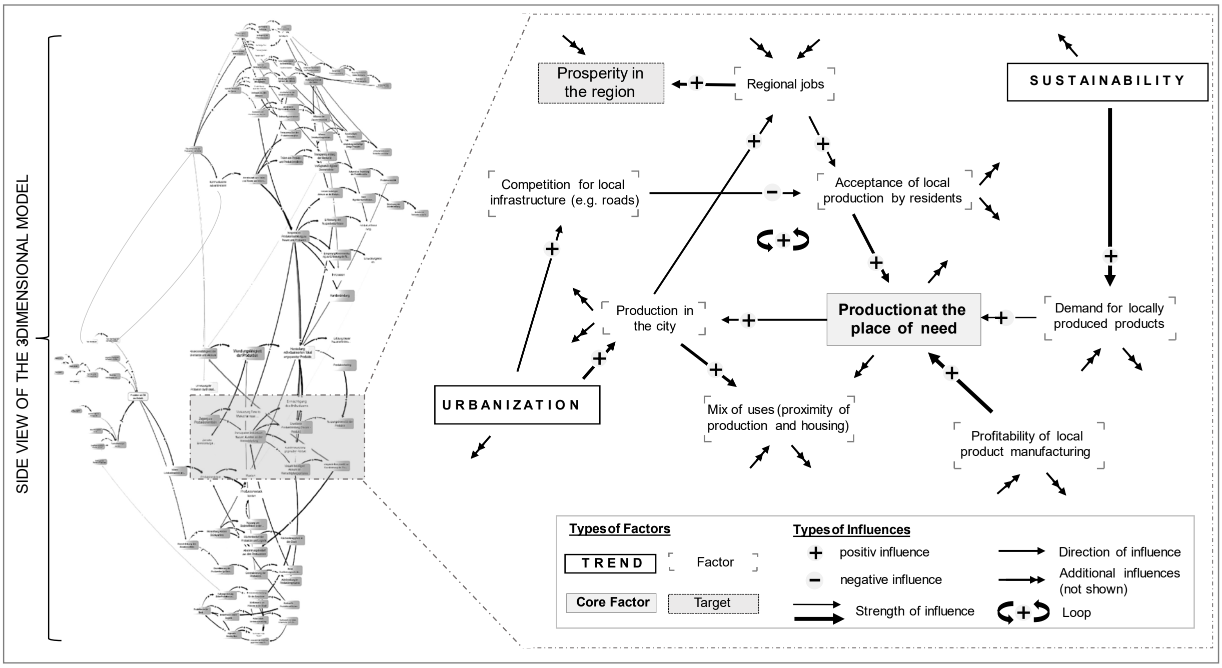


Figure 1: Model of Local Production with an example close-up of an effect chain, showing production at the place of need as the central factor and selected direct and indirect influences

4.1 Key factors influencing the local production of goods

The dimension of **local production of goods** describes the spatially concentrated production at the place of demand. In the model, this dimension is represented by the factor ***production at the place of need***. Driven by the trend of *urbanization* [16] *production at the place of need* will increasingly become *production in the city* in the future.

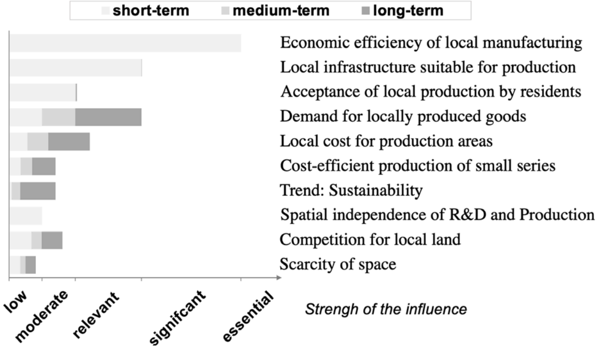


Diagram 1: Top 10 factors influencing production at the place of need

Diagram 1 shows the most important influences on the factor ***production at the place of need*** according to the model evaluation. The influences are intentionally not marked as positive or negative, since the manifestation of the factor determines that, e.g. *if the local cost for production areas* is low, the influence on ***production at the place of need*** is positive, but if the cost is high, it is negative.

Production at the place of need stimulates *sustainable product manufacturing* by avoiding transport [17,18]. It reduces *risks and costs of global logistics* [19,4] and strengthens regional value creation structures and thus provides *prosperity in the region*. Furthermore, it has a direct impact on *high quality of local life* [20,21]. This factor is significantly determined by the *economic efficiency of local product manufacturing*. By reducing the sales market to the local area, *cost-efficient production of small series* [22] and the *spatial concentration of demand* (through urbanization) will have an increasingly, reinforcing influence on ***production at the place of need*** [23]. In contrast, *rising costs for land in urban areas* (trend: *urbanization*) [24,20,25,26] will negatively influence *economic efficiency of local product manufacturing* in the future. Current trends toward more *sustainable consumption*, which will intensify in the future, will promote the emergence of a ***production at the place of need*** by increasing *demand for locally produced products* [3,4]. The availability of a *local infrastructure suitable for production* [24,21] and the *acceptance of local residents [27]* towards these forms of production remain essential for the development of local manufacturing [27].

However, some influences of ***production at the place of need*** mentioned in the current academic discourse could not be confirmed within the model. I.e. *urbanization*, which has an ambivalent influence on production on site. On the one hand it promotes the *emergence of local agglomerations* and thus the *spatial concentration of demand*. On the other hand, it increases *competition for local land* [28,25], which increases the *cost of local production space* [20] and reduces the *economic efficiency of local manufacturing.* Furthermore, the increasing *risks and costs of global logistic* do not significantly influence ***production at the place of need***, as they only affect *the economic efficiency of local product manufacturing* to a small to moderate extent.

4.2 Key factors influencing the utilization of local resources

The **use of local resources** in a local production encompasses the ***use of local (raw) materials*** and the involvement of regional actors, companies as well as workers, represented in the model by the factor ***implementation of production by local stakeholders***.

Use of local (raw) materials and ***implementation of production by local stakeholders*** strengthen the local value creation, support the creation and retention of *regional jobs* [3,29,17] and ultimately increase the *prosperity in the region* [20]. Additionally, the ***use of local (raw) materials*** through the downsizing of value-added cycles and *reduction of the global transport of goods* leads to *sustainable product creation* [4,23].

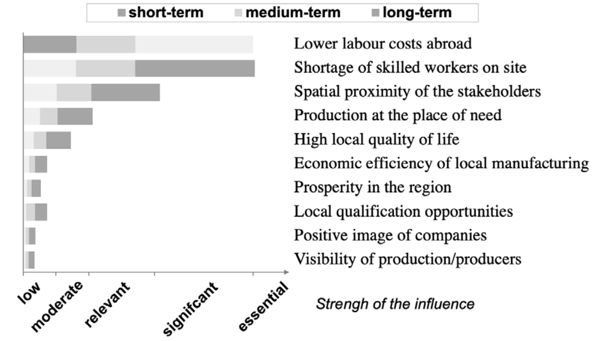


Diagram 2: Top 10 factors influencing the implementation of production by local stakeholders

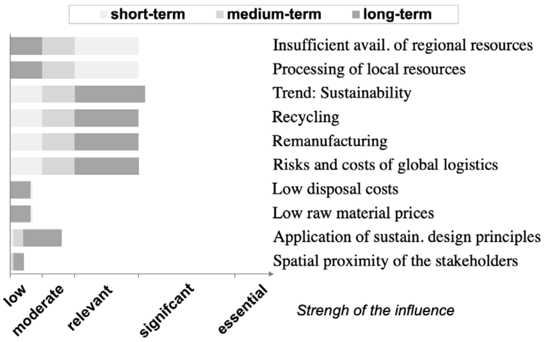


Diagram 3: Top 10 factors influencing use of local (raw) materials

Diagram 2 and Diagram 3 show the comparative influence of the key factors on ***implementation of production by local stakeholders*** and ***use of local (raw) materials*** respectively.

At the moment, the ***use of local (raw) materials*** in local value-added cycles is especially determined by the *(insufficient) availability of regional resources* [30,22,1] and by *possibilities for the economic extraction and processing of local resources.* The relevance of local resource extraction, however, will decrease in the future while the recirculation of materials and products into value-added cycles will be raised. The increase of currently *low disposal costs* and *low raw material prices* as well as the consistent *application of sustainable design principles* [30], the *modularity of products* [31,4], the enhanced *transparency along the value chain* [32] and *along the product life cycle* will mid- to long-term significantly boost the relevance of *recycling* as well as *remanufacturing* for the ***use of local (raw) materials*** in the context of local value creation.

While the *spatial proximity of the stakeholders (producer and user)* does promote *recycling* as well as *remanufacturing* [33], it holds only secondary relevance compared to the aforementioned drivers (refer to

Diagram 3). The *trend: sustainability* and the related change of values will influence consumer behaviour long-term in so far that consumers will specifically call for the use and re-use of local resources [3].

The ***implementation of production by local stakeholders*** is dominated by the mostly *lower labour costs abroad* [4,34]. This factor's relevance will however lessen in the future due to the assimilation of labour costs and due to attempts to impose standards along the value chain through regulations (e.g. supply chain law). The *shortage of skilled workers on site* will determine the ***implementation of production by local stakeholders*** mid- to long-term [35,27]. The involvement of local stakeholders will additionally be impacted by the ***production at the place of need*** and the related *spatial proximity of the stakeholders* [3,20,27].

4.3 Key factors influencing the addressing of local demands

The dimension **addressing of local needs** is represented in the model by the following factors: ***on demand production*** and ***production of individualised / locally adapted products***. The ***production of individualised / locally adapted products*** refers to the potential of local manufacturing to adapt to the local or individual requirements of regional users (e.g. in terms of function, design). ***On demand production*** represents the ability to respond to local demand quickly and in the required quantities.

Responding to local demands promotes *sustainable manufacturing* by avoiding overproduction and warehousing through ad hoc demand-driven production (on-demand) [28]. In addition, local, specific user requirements are fulfilled and customer acceptance of locally manufactured products is increased [4,32].

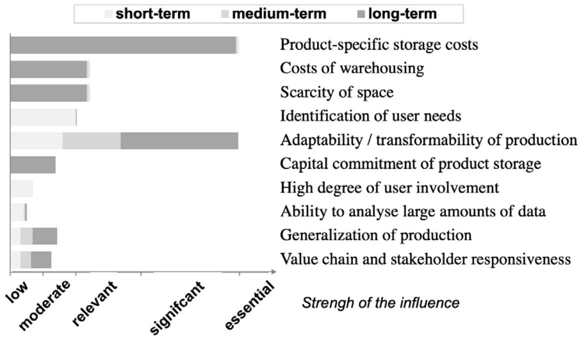


Diagram 4: Top 10 factors influencing on demand production

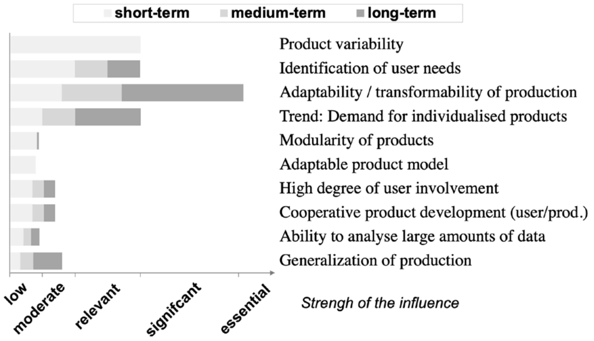


Diagram 5: Top 10 factors influencing production of individualised products

Diagram 4 and Diagram 5 show the compared relevance of the main factors to ***on demand production*** and ***production of individualised / locally adapted products*** in the context of a local production over time.

The *adaptability* of local production determines its potential for the targeted **fulfilment of local needs** [30]. In the future, the importance of this factor will increase due to *new, highly flexible production technologies* [4,22,34], the *generalization of production* and the increased use of *modular product structures* [31]. The *variability* of locally produced goods also promotes **addressing local needs,** for instance through functional variability (e.g. through software customisation, modular product design) and product scale variability (e.g. the possibility to choose the production process depending on the production quantities) [34]. Knowledge about local needs is gained through targeted *identification of user needs*, which has a high impact on **addressing the local needs** by concepts like *co-creation or co-design* [19,30,18] and the increasing *ability to analyse large amounts of data* (data mining, artificial intelligence) [34].

In the future, changing consumer demands for *individualisation* and *immediate availability of products* will promote the emergence of *demand-driven, ad hoc production* in local contexts.

The *scarcity of space* and related *storage costs* also have an important influence on ***on-demand production*** [18] in systems of local value creation when located in urban areas.

In contrast to some views in the literature [30,4,18], our model did not show that *spatial proximity between producers and consumers* facilitates **addressing local needs**. Instead, it can be assumed that knowledge of local needs can primarily be generated in ways other than spatial proximity (e.g. virtual cooperation between customer and producer in co-creation formats).

4.4 Summary of the findings

For an easier comparison of the analysis' results, the most important factors of local manufacturing, the primary target factors and the key factors are summarized in Table 2.

Table 2: Summary of the main attributes, target factors and key influencing factors of local manufacturing

Dimension of local manufacturing	**(A) Local production of goods**	**(b) Utilisation of local resources**		**(c) Addressing of local demands**	
Main attributes of local production	Production at the place of need	Use of local raw materials and materials	Implementation of production by local stakeholders	Production of individualised / locally adapted products	On demand production
Primary target factors	o Sustainable production of goods o Prosperity in the region	o Sustainable production of goods o Prosperity in the region	o Prosperity in the region	o Sustainable production of goods o Fulfilment of local consumer demands	o Sustainable production of goods o Fulfilment of local consumer demands
Key influencing factors	o Economic efficiency of local manufacturing o Local infrastructure suitable for production o Demand for locally produced goods o Acceptance of local production by residents o Local costs for production areas o Cost-efficient production of small series o Trend: Sustainability o Spatial independence of R&D and Production o Competition for local land o Scarcity of space	o Recycling o Re-Manufacturing o Risks and Costs of global logistics o Trend: Sustainability o Availability of regional resources o Processing of local resources o Application of sustainable design principles o Disposal costs o Prices of raw material	o Lower labour costs abroad o Shortage of skilled workers on site o Spacial proximity of the stakeholders o Production at the place of need o High quality of life on site o Economic efficiency of local manufacturing	o Adaptability / transformability of production o Identification of users need o Product variability o Trend: Demand for individualised products o Generalization of production o High degree of user involvement o Cooperative product development o Modularity of products o Adaptable product model o Ability to analyse large amounts of data	o Product-specific storage costs o Adaptability / transformability of production o Cost of warehousing o Scarcity of space o Scarcity of land o Identification of users need o Generalization of production o Value chain and stakeholder responsiveness o Capital commitment of product storage o High degree of user involvement o Ability to analyse large amounts of data

The key factors influencing local manufacturing can be differentiated by being based primarily on technological, economic, political or societal developments.

Product manufacturing **addressing local needs** is mainly driven by technological developments that focus on the adaptability of local value creation systems and the recording of the user's needs.

The development of **production at the place of need** is primarily influenced by political and societal drivers. While the adaptability of value creation systems does have a relevant influence on the economic production on-site, the negative influence of rapidly rising costs for local production sites caused by the merging of conurbations is more important. This challenge cannot only be solved technologically (e.g. by the downsizing and adapting of production technologies), but through political regulation. The availability of a suitable infrastructure for local manufacturing is also dependent on political decisions. Relevant societal drivers are the changing consumer behaviour and the rising demand for locally produced goods.

The **use of local (raw) materials** benefits on a technological level from material and process innovations in order to achieve effective and efficient recycling and remanufacturing processes. The central drivers are rising costs for raw materials, energy and waste disposal (economic drivers), which depend on political decisions (e.g. CO_2 taxes, export bans on plastic waste). The implementation of production by local stakeholders can benefit from the growing assimilation of labour costs as well as the political measures to avert the worsening shortage of skilled workers on-site.

The differentiation between the key drivers of the main attribute of local manufacturing shows, that forms of sustainable, local value creation are not primarily driven by technological, operational or business model innovations, but by a combination of political and societal developments. In the end, political decisions will determine a production at the place of need (availability of space and of a suitable infrastructure) by local actors (international assimilation of labour costs, aversion of a shortage of skilled workers) while using local resources (promotion of a regional circular economy) for the fulfilment of local demand.

In order to base such political decisions to promote local manufacturing on scientific findings, further research is needed on the actual potentials and implementations for local production systems. In particular, multidimensional benchmarks must be developed in comparison to a global, industrial value creation.

5. Discussion

Although the model shows the strengths of the influences of the mentioned factors in comparison to one another and describes potential short-, mid- and long-term developments of these factors, is has some limitations. With its qualitative nature it sharpens the understanding for the systematics of local manufacturing, but it does not say anything about the probability of the actual occurrence of concrete developments. Therefore, quantitative prognostic methods are needed. The complexity of the causal effects of the object of study, however, would not allow for a consistent quantification, which is why a qualitative approach was chosen here. In a next step, the model could be expanded through a systematic analysis of especially relevant subsystems, which are aimed at the quantification of the causal effects in order to be able to make more concrete statements regarding the development of the system within defined borders.

6. Outlook

The presented model is able to sharpen the understanding for the key factors of local manufacturing on a technological, societal, political and economic level. Using these findings various scenarios can be generated to show different paths of development for local production systems. From these options concrete recommendations of action for political, societal, operational and technological stakeholders can be derived in order to reach the goals mentioned in the introduction of this paper (e.g. reduction of CO_2 emissions, empowerment of local actors). Nevertheless, the model cannot be considered complete since factors such as business taxes, etc. have not yet been depicted and would have to be added in the course of further text analyses.

Acknowledgements

This research is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr [project: Production Next Door].

References

[1] Srai, J.S., Harrington, T.S., Tiwari, M.K., 2016. Characteristics of redistributed manufacturing systems: a comparative study of emerging industry supply networks. International Journal of Production Research 54 (23), 6936–6955.

[2] Bortolini, M., Galizia, F.G., Mora, C., 2018. Reconfigurable manufacturing systems: Literature review and research trend. Journal of Manufacturing Systems 49, 93–106.

[3] Dieter Läpple, 2013. Produktion zurück in die Stadt?, in: Kronauer, M., Siebel, W. (Eds.), Polarisierte Städte. Soziale Ungleichheit als Herausforderung für die Stadtpolitik. Campus-Verl., Frankfurt am Main, 130–150.

[4] Matt, D.T., Rauch, E., Dallasega, P., 2015. Trends towards Distributed Manufacturing Systems and Modern Forms for their Design. Procedia CIRP 33, 185–190.

[5] Ellen MacArthur Stiftung, 2013. Towards the Circular Economy.

[6] Krenz, P., Stoltenberg, L., Markert, J., Saubke, D., Redlich, T., 2022. The Phenomenon of Local Manufacturing: An Attempt at a Differentiation of Distributed, Re-distributed and Urban Manufacturing, in: Andersen, A.-L., Andersen et al. (Eds.), Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems (CARV2021/MCPC2021), Aalborg, Denmark, 1st ed. 2022 ed. Springer International Publishing, 1014–1022.

[7] Grossmann, C., 1992. Komplexitätsbewältigung im Management: Anleitungen, integrierte Methodik und Anwendungsbeispiele. Zugl.: St. Gallen, Hochsch., Diss.: 1992. Verl. GNC, Winterthur.

[8] Vester, F., 1991. Ausfahrt Zukunft Supplement: Material zur Systemuntersuchung, München.

[9] Vester, F., 2019. Die Kunst vernetzt zu denken: Ideen und Werkzeuge für einen neuen Umgang mit Komplexität. Ein Bericht an den Club of Rome. Deutsche Verlagsanstalt GmbH, München.

[10] Vester, F., Hesler, A. von, 1980. Sensitivitätsmodell.

[11] Vahs, D., 2005. Organisation: Einführung in die Organisationstheorie und -praxis, 5., überarb. Aufl. ed. Schäffer-Poeschel, Stuttgart.

[12] Redlich, T., Wulfsberg, J.P., 2011. Wertschöpfung in der Bottom-up-Ökonomie. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg.

[13] Beer, S., 1970. Kybernetik und Management, 4. Aufl., 19.- 21. Tsd ed. Fischer, Frankfurt/M.,

[14] Krenz, P., 2020. Formen der Wissensarbeit in einer vernetzten Wertschöpfung: Formen der Wissensarbeit in einer vernetzten Wertschöpfung. Helmut-Schmidt-Universität.

[15] Strübing, J., 2014. Grounded Theory: Zur sozialtheoretischen und epistemologischen Fundierung eines pragmatistischen Forschungsstils, 3., überarb. u. erw. Aufl. ed. Springer VS, Wiesbaden.

[16] Burggräf, P. et al., 2019. Urban factories: Industry insights and empirical evidence within manufacturing companies in German-speaking countries. Procedia Manufacturing 28, 83–89.

[17] Gärtner, S., Stegmann, T., 2015. Neue Arbeit und Produktion im Quartier: Beobachtungen und Wishful Thinking. Forschung Aktuell 7/2015. Gelsenkirchen: Institut Arbeit und Technik (IAT).

[18] Petrulaityte, A., Ceschin, F., Pei, E., Harrison, D., 2017. Supporting Sustainable Product-Service System Implementation through Distributed Manufacturing. Procedia CIRP 64, 375–380.

[19] Kohtala, C., 2015. Addressing sustainability in research on distributed production: an integrated literature review. Journal of Cleaner Production (106), 654–668.

[20] Lentes, J., 2016. Urbane Produktion, in: Spath, D., Westkämper, E. (Eds.), Handbuch Unternehmensorganisation. Strategien, Planung, Umsetzung, Living reference work, continuously updated edition ed. Springer Vieweg, Berlin, Heidelberg, 1–11.

[21] Mistry, N., Byron, J., 2011. The Federal Role in Supporting Urban Manufacturing, April 29.

[22] Pearson H, Noble G, Hawkins J., 2013. Re-distributed manufacturing workshop report.

[23] Moreno, M., Turner, C., Tiwari, A., Hutabarat, W., Charnley, F., Widjaja, D., Mondini, L., 2017. Re-distributed Manufacturing to Achieve a Circular Economy: A Case Study Utilizing IDEF0 Modeling. Procedia CIRP 63, 686–691.

[24] Erbstößer, A.-C., 2016. Produktion in der Stadt: Berliner Mischung 2.0.

[25] Sassen, S., 2009. Cities Today: A New Frontier for Major Developments. The Annals of the American Academy of Political and Social Science 626 (1), 53–71.

[26] Schrock, G., Wolf-Powers, L., 2019. Opportunities and risks of localised industrial policy: the case of "maker-entrepreneurial ecosystems" in the USA. Cambridge Journal of Regions, Economy and Society 12 (3), 369–384.

[27] Schössler, M., et al., 2012. Future Urban Industries - Produktion, Industrie, Stadtzukunft, Wachstum. Wie können wir den Herausforderungen begegnen? Policy Brief 11/12. Stiftung neue Verantwortung.

[28] Fuchs, M. et al., 2017. Urbane Produktion: Dynamisierung stadtregionaler Arbeitsmärkte durch Digitalisierung und Industrie 4.0? Working Paper, Köln, Aachen.

[29] Freeman, R., McMahon, C., Godfrey, P., 2016. Design of an Integrated Assessment of Re-distributed Manufacturing for the Sustainable, Resilient City, in: Setchi, R., Howlett, R.J., Liu, Y., Theobald, P. (Eds.), Sustainable Design and Manufacturing 2016. Springer Int. Publ., 601–612.

[30] Lowe, A.S., 2019. Distributed Manufacturing: Make Things Where You Need Them, in: Redlich, T., Wulfsberg, J.P., Moritz, M. (Eds.), Co-creation. Reshaping business and society in the era of bottom-up Economics. Springer, Cham, Switzerland, 37–50.

[31] Lentes, J., Hertwig, M., Zimmermann, N., Mahlau, L.-M., 2018. Development Path for Industrrial Enterprises towards Urban Manufacturing. dtetr (icpr).

[32] Yakovleva, N., Frei, R., Rama Murthy, S., 2019. Sustainable Development Goals and Sustainable Supply Chains in the Post-global Economy. Springer International Publishing, Cham.

[33] Turner, C., Moreno, M., Mondini, L., Salonitis, K., Charnley, F., Tiwari, A., Hutabarat, W., 2019. Sustainable Production in a Circular Economy: A Business Model for Re-Distributed Manufacturing. Sustainability 11 (16).

[34] The Government Office for Science, London, 2013. Foresight (2013). The Future of Manufacturing: A new era of opportunity and challenge for the UK. Summary Report.

[35] Busch, H.-C., Mühl, C., Fuchs, M., Fromhold-Eisebith, M., 2020. Hybride Formen urbaner Produktion durch Digitalisierung? Trends und Beispiele aus Nordrhein-Westfalen. Raumforschung und Raumordnung Spatial Research and Planning 78 (4), 321–336.

Biography

Pascal Krenz, Dr.-Ing. (*1983) has been a research assistant at the Institute of Production Engineering (LaFT) within the Department of Mechanical Engineering of Helmut-Schmidt-University in Hamburg (HSU HH) since 2011. He leads the research area on local production at the Institute and related research projects.

Dominik Saubke, M.Sc. (*1989) has been a research assistant at the Institute of Production Engineering (LaFT) of HSU HH since 2021. With several years of experience in Product Engineering and Product Development his current research focus is production-readiness in collaborative crowd design processes.

Lisa Stoltenberg, M.A. (*1991) has been a research assistant at the Institute of Production Engineering (LaFT) at HSU HH since 2021. She studied social sciences and is currently working on topics in the field of consumer research, with a focus on sustainable purchasing decisions.

Julia Markert, M.Sc. (*1996) has been a research assistant at the Institute of Production Engineering (LaFT) of HSU HH since 2021. She has a background in Production Engineering; her current research focus is on the creation and coordination of dynamic value chains in local production networks.

Tobias Redlich, Dr.-Ing. (*1981) has been chief engineer at the Institute of Production Engineering (LaFT) within the Department of Mechanical Engineering of Helmut-Schmidt-University in Hamburg (HSU HH) since 2013.

3rd Conference on Production Systems and Logistics

Local Differential Privacy In Smart Manufacturing: Application Scenario, Mechanisms and Tools

Sascha Gärtner[1], Michael Oberle[1]
[1]*Fraunhofer IPA, Stuttgart, Germany*

Abstract

To utilize the potential of machine and deep learning, enormous amounts of data are required. A common and beneficial approach is to share datasets between the parties involved for training purposes or even to release datasets to the public. However, several incidents have shown that despite anonymizing the data, attackers are still capable of identifying individuals in the data and extracting their sensitive information. The methods of differential privacy address this problem by adding a statistical noise to data points in the shared dataset. Since manufacturing data not only contains information about individual persons but also about the companies, their process knowledge, products, and orders add more complexity to the application of differential privacy compared to other domains. In this paper, we highlight why conventional methods of anonymization are not sufficient to guarantee data protection and thus present the necessity of using differential privacy. To illustrate its usefulness for manufacturing we present a specifc application scenario and examine potential threats when sharing manufacturing data. We identify mechanisms to perturbate data and map these to variable types in the manufacturing context. To guide practical application and research we finally outline existing differntial privacy libraries, and highlight current limitations.

Keywords

Local differential privacy; smart manufacturing, industrial Internet of Things, privacy preservation, LDP;

1. Introduction

Driven by the digital transformation, traditional manufacturing is currently undergoing a change towards smart manufacturing. The digital transformation is especially initiated by key technologies such as the Internet of Things, 5G, CPSs, BigData, and Artificial Intelligence (AI) [1,2]. Due to the application of sensor technology and IT infrastructure, the amount of data generated during manufacturing processes are constantly increasing [3]. However, the amounts of data are often not sufficient to train generalizable machine learning or AI models. To address this issue, approaches such as the establishment of common data spaces and federated learning for collaborative training of algorithms are emerging [4]. Numerous real-world examples [5–7] have shown that sharing data with third parties is critical in terms of privacy violations. Typically, privacy is violated when attackers identify individuals in the published dataset and thus have access to sensitive information [8]. In the context of manufacturing, privacy threats are much more complex and increase since additional sensitive company-relevant data can be identified. For this reason, we contribute by mapping the concepts of differential privacy to the manufacturing context and presenting a real-world application scenario including perturbation mechanisms. At the beginning of the paper, we give an overview of the general motivations of differential privacy by comparing conventional methods and mentioning well-known privacy leaks. We then map the concept of differential privacy to manufacturing and specify the problem to machine manufacturers and their customers who operate the machines in their

DOI: https://doi.org/10.15488/12125

factories. We then identify relevant parameters that occur in manufacturing, categorize them by variable type, and present examples of suitable differentially private mechanisms. Finally, we provide an overview of libraries and toolboxes for guiding the practical application of differential privacy, outline topics for further research and conclude the paper by highlighting the key points.

1.1 Privacy concerns and the need for differential privacy

Typical techniques to ensure the privacy of user data are masking, generalization, and k-anonymization. For additional security, these techniques can be complemented by encryption mechanisms such as homormorphic encryption [9]. However, the use of these techniques is vulnerable to a variety of attacks (e.g., linkage, reconstruction, and differentiation attacks). Linkage attacks for example use similar publicly available datasets to find similarities within the data. It has been proven that even a few data points are sufficient to uniquely identify individual persons [10,7]. The encryption of individual sensitive data points in a data series also involves vulnerabilities. If the attacker succeeds in gaining knowledge of the function used for encryption, the data can be decrypted again by systematically testing possible input values [11,12].

To overcome these problems, the research field differential privacy emerged. Differential privacy can be seen as a process A, applied to some data D. The process might be the estimation of the mean over the distribution of a dataset or a machine learning process to predict values. To achieve the formal definition of differential privacy the process A has to be modified. This is usually done by adding noise at a certain point in the process. Adding the right amount of noise strongly depends on the use case and threat model.

Considering two neighboring datasets $D_1$and D_2, where dataset D_2 differs from dataset D_1 by just a record of a person, the process A is considered ε-differentially private if the output O of the process is approximately the same when being applied to both datasets. This leads to approximately identical probabilities $\mathbb{P}$. The relationship between the two probabilities is described by the following definition [13].

$$\mathbb{P}[A(D_1) = O] \leq e^{\varepsilon} \cdot \mathbb{P}[A(D_2) = O] \tag{1}$$

The mechanism used to add noise is dependent on the data type. Typically, the Randomized Response, Laplace, Gaussian, and Exponential mechanisms are used. The parameterization must be adapted in each case to the variable to be determined. Several real-world applications demonstrate the potential, but also the complexity, of differential privacy. Apple uses differential privacy to collect data from end-users of iOS or macOS [14,15]. For example, words that are typed by a sufficient number of users but are not yet in the dictionary are collected differentially private. Facebook created and released a dataset that provides information about user interactions with websites that have been shared on their platform [16].

2. Scenario of differential privacy in manufacturing

While the purpose of Differential privacy is easily accessible when exposing data to the general public to protect the privacy of individuals, the transfer of use cases to manufacturing is not immediately apparent. Considering the paradigm shift from traditional production to autonomous manufacturing, the relevance of data-driven approaches to make manufacturing processes more efficient is constantly increasing. The importance of data for the optimization of processes, plants, and machines is accordingly high. Sharing company-related or process data in manufacturing is therefore unavoidable for companies. [17]

During our research, we identified a common scenario that represents the current issues and concerns of manufacturing companies in merging and sharing data for training machine learning algorithms. There is a trend to improve the customer's process by offering value-added services additional to the machine itself, thus opening up new business areas [18]. The machine manufacturer (curator), wants to collect data from the machines in the customer's productive operation in order to subsequently optimize the machine. The motivations can be constructive improvements of the machine through insights into the daily production operation, quality checks or improved control loops of the machine. [19]

Regardless of whether the machine manufacturer wants to process the aggregated data of the customer with statistical methods, machine learning, or artificial intelligence, there are two possible ways (Figure 1) for the customer to share his data. The differential private mechanism $M$is either held by the curator (GDP-Model) or by the customer (LDP).[20]

In the global differential privacy (GDP) model, the customer can share his raw data with the curator. In this case, the customer has to fully trust the curator. It is not necessarily defined how and whether the curator provides the data to third parties or other customers. The curator can aggregate the datasets of individual customers and thus host a dataset in total. Other customers or external parties can make requests to learn distributions of certain quantities in the dataset. The privacy of the customer data can be protected if the output of the query is appropriately noisy.

In the local differential privacy (LDP) model, noise is added to raw data before sending it to the curator. This model has the advantage that the curator does not have to be trusted. It should be noted that data can be of any variable type. The concept of federated learning enables distributed learning of a shared neural network [4]. In this case, the curator only aggregates the weights of the model. Prior to this, the client trains the model on a local instance (e.g. edge-device). Privacy is achieved by adding noise to the gradients, objective, or output during the training of the model. The biggest advantage of federated learning is that the mechanism works reliably regardless of the data type. However, there are several disadvantages. The customer must have the computing resources to perform the training of the neural network on the edge. In addition, qualified specialists are needed to implement the necessary IT infrastructure and pipelines [21]. In consequence this leads to high costs related to setup and operation. If the purpose and benefits are not immediately apparent to the customer, skepticism arises and they are not willing to make investments. For this reason, we refer to a scenario in which the customer has no computing resources locally available.

A more effective solution in terms of cost and effort is offered by adding noise directly to the raw data. Due to many different types of parameters and attributes occurring on the shopfloor in manufacturing, adding noise to raw data is more complex and must be adapted to the individual case. To select suitable mechanisms and therefore ensure privacy protection it is necessary to identify type of the variable first.

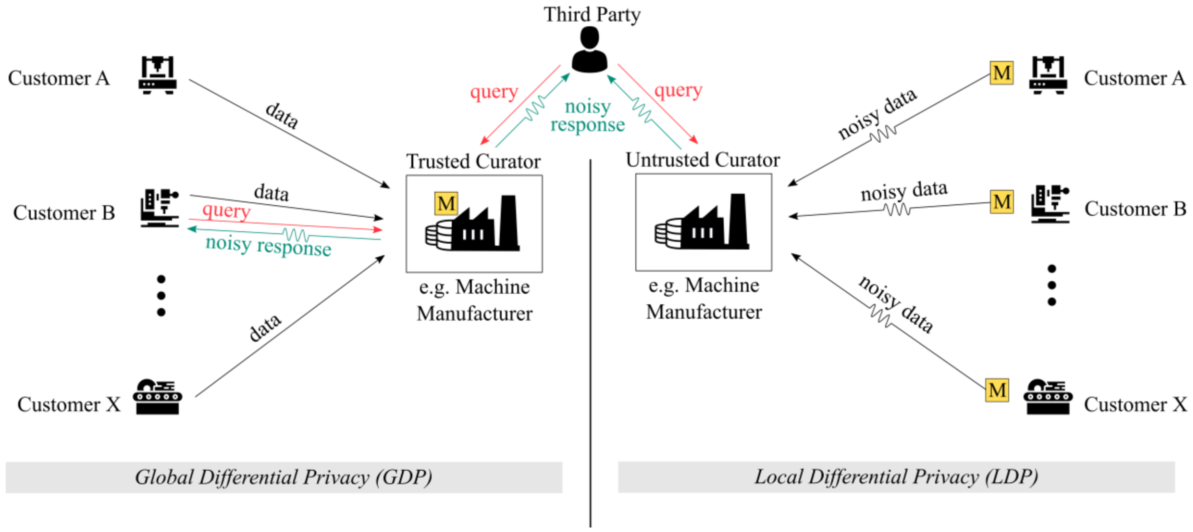


Figure 1: Comparison between global and local differential privacy

3. Parameters in manufacturing and corresponding threats

In order to apply differential privacy to use cases in the manufacturing context, we first determine parameters that occur in the manufacturing environment. The parameters can be divided into different categories and specified (see Table 1). Each parameter category represents different types of information. Consequently, there are different risks in the case of a violation of privacy. While the manufacturing process parameters describe the manufacturing process itself, the process parameters further specify the process with its attributes. These parameters contain the manufacturer's main process knowledge for creating the product. The environment condition parameters have an indirect influence on the process. The ambient temperature can influence the thermal expansion of components within the machine and the workpiece itself, which in turn affects the quality of the workpiece. The attributes listed below can be of different variable types. For the application of mechanisms to satisfy differential privacy, the variable type plays a decisive role.

Table 1: Different kinds of sensitive parameters occurring in the manufacturing context according to [22]

Parameter Category	**Attribute**
Manufacturing Process Parameters	milling, turning, laser cutting, welding, casting, extrusion, stamping, assembling, etc.
Process Parameters	spindle speed, cutting speed, pressure, coolant, voltage, feed, current, force, torque, etc.
Environment Condition Parameters	humidity, temperature, date, time, rainfall, etc.
Working Condition Parameters	duration, shift, worker id, machine id, etc.
Target Parameters	quality, yield, productivity, OEE, KPIs, etc.
Other Parameters	manufacturing order, material, production numbers, geometric data, position, etc.

If we assume a company shares a large dataset containing the variables listed above multiple threats about the company’s process knowledge as well as sensitive business information can arise which are not immediately apparent. A potential attacker can analyze environment condition parameters in the dataset and combine them with publicly available weather data. If there are sufficient matches over time it is possible to identify the company itself at its location. Together with other parameters such as manufacturing orders, date, material, supply chains, or even suppliers thus the order situation might be revealed. If the data is then complemented with target parameters, attackers can reveal the company's productivity and efficiency in production. Therefore, it would be possible to estimate the turnover or profit of a company.

Corporate-related threats link to the process knowledge, which is required to produce workpieces efficiently while ensuring quality. Potential attackers can use the type of the manufacturing process, machine ID, and the associated process parameters to gain sensitive process knowledge about the setup of the machine during the ramp-up process. Another threat can occur if the position data of AGVs is regarded in conjunction with the time stamp. This allows attackers to determine the factory layout and retrace the production routes within the factory. The previously mentioned threats should not be regarded as comprehensive. These are only a few examples for illustrating the problems that can arise from exposing the mentioned variables.

4. Data perturbation mechanisms

Different mechanisms can be used to prevent the exposure of the previously listed threats by adding noise to data before sharing. However, the mechanism used is directly dependent on the variable type and its parameterization is not trivial. In the following, the most common variable types are explained and their context for manufacturing is presented. In addition, the mechanism by which differential privacy can be satisfied and the analogies to related areas are described.

Essentially, the Randomized Response, Laplace Mechanism, Gaussian Mechanism, or Exponential Mechanism are used to add noise [23,24]. Depending on the application case the functions are parameterized by the so-called sensitivity. Hereby the privacy budget from low to high can be adjusted.

Most publications and thus algorithms refer to the assumption that a user owns a data set and a potential attacker issues a query (e.g. mean value of a numeric variable) to the user's database. The query can be connected by AND, OR conditions [25]. In this case, the noise will be applied to the true value and then sent to the potential attacker. Depending on the sensitivity, the database size, and the number of variables, the number of possible queries of a user has to be limited. In addition, the number of combinations of AND, OR conditions also has to be limited. Since the determination of the noise, as well as the boundary conditions to the queries, must be individually adjusted for each problem, the implementation is very extensive.

Our desired approach is an algorithm that allows publishing the dataset under conditions of differential privacy instead of hosting a dataset and answering the queries of clients. The goal is to apply noise to a dataset so that it can be passed on to third parties without being concerned about query limitations. Sensitive information or process variables can not be identified in this case. But the data should still contain enough information to learn the statistical distribution of the variable. For example, it would still be possible to learn something about the wear of the machine and upcoming maintenance but it is not possible to determine the exact process parameters used for manufacturing the workpieces. In reality, many different and complex data structures exist. In the following, we present suitable mechanisms for each variable type and highlight the relation to manufacturing.

4.1 Binary data

The most straightforward approach to publishing a variable differentially private is for binary variables. A very effective approach is the so-called Randomized Response.[26] As an example, the question is asked whether the machine owner (user) has used coolant in a certain process section. The possible answers in this scenario are just *yes=1 or no=0*. Before answering the question, an imaginary coin is tossed. If the coin lands on heads, the user must answer truthfully. If the coin lands on tails, a second coin is tossed. If the second coin lands on heads the user answer with *yes*, otherwise with *no*. This algorithm is differentially private by definition. Therefore, the user can safely reveal the truth. Though noise is also being added to the data by the mechanism. However, if a large set of responses from different companies is received, the noise can be canceled out and the statistical distribution of the use of coolant in the process can be determined. However, it cannot be determined whether a specific company used coolant in its process or not. The most known application of this algorithm is called RAPPOR (extended with additional operations). It was developed and used by Google to determine the default search engines of Google Chrome users [27].

4.2 Numerical data

Applying noise to numerical data can be achieved by using the Laplace or Gaussian mechanism [28]. These mechanisms represent a distribution with probability values and a scaling factor. If we ask a machine user how often the fixture did break during the last year, we expect a numerical value as a response. Before publishing the true value, the dataset holder selects a random value from the Laplace distribution, adds it to the true value, and then writes the perturbed value to the dataset. Instead of the correct number of fixture breakages, a noisy value is given. If multiple machine users report their insert breakage data, the machine manufacturer will be able to determine the average without revealing how often the insert really broke at each user. Other numerical queries could be the calculation of the mean over several data rows or the query about the numerical distribution of a parameter. For example, a querier (machine manufacturer) may ask the question, how often machine failures in the range of [0-9;10-19;20-29;…] occurred. The output would thus be a histogram, which can also be released under the satisfaction of differential privacy. It is possible to apply an individual noise to each count by applying the distribution function to each single count.

4.3 Categorical data

Publishing categorical variables taking into account differential privacy can be seen as an extended version of Randomized Response. But instead of two categories (0 and 1), finite categories are possible. An implementation is the Google algorithm RAPPOR [29]. Other algorithms are the Local Hash method [30] and the Unary Encoding Method [31] which is the basic concept of RAPPOR. The Unary Encoding method (Figure 2) is very intuitive and is presented using an example in the manufacturing context. A querier (machine manufacturer) wants to know which clamping tools his customers use in their production. All parties agree that there are four different possibilities in total. These four possibilities are each represented by a position within a bit string. Position 1 in the bit string stands for the three-jaw chuck, position 2 for the four-jaw chuck, position 3 for the collet chuck, and position 4 for the centering tip. The customer now encodes his clamping tool used in production into the bit string. Then each bit is perturbed according to the Randomized Response method. The perturbed bit strings will then be sent to the querier. The querier adds up the individual positions in the bit string and can thus calculate a distribution of the clamping tools used. However, the querier does not know which exact clamping tool is used by which customer. It should be noted that the accuracy increases significantly with a higher number of contributions.

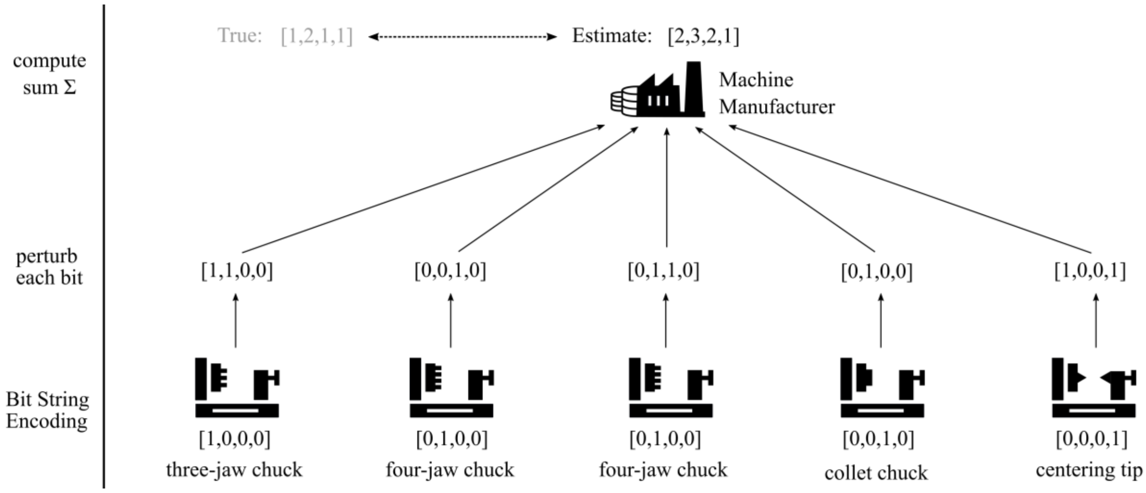


Figure 2: Example of Unary Encoding according to [31]

4.4 Time series data

Time series are becoming increasingly important due to smart sensor technology and advances in data transmission rates. It is possible to reliably record data at high sampling rates, to process and store it outside the PLC. In the context of data analysis for monitoring machines or predicting the occurrence of events, the analysis of time series is of great relevance. In addition to sensor data, time series can also be multidimensional position data from IoT devices on the shopfloor. There is no generally valid way to obtain information from time series. For example, a time series (e.g. sensor signal) can be sampled with sliding windows. A querier might be interested in the average value of each sliding window. Then response would just be a noisy numeric value, as seen above. The analogy to the histogram mentioned before would be if the querier asks for the frequency spectrum of a periodic signal. However, it is also possible to publish several time series under the condition of differential privacy. For example, this can be achieved by a re-quantization mapping [32]. Another reference introduces a perturbation mechanism consisting of several single steps to satisfy differential privacy [33].

Coordinate, position, or trajectory data often differ from one-dimensional time series by the property of multidimensionality. Publishing coordinate data is rarely addressed in the literature. However, some papers show possible ways how this can be achieved. [34–36]

5. Available libraries and tools for implementing differential privacy

In order to apply differential privacy efficiently in manufacturing scenarios, it is necessary to keep the implementation as simple as possible. It should be avoided to implement the algorithms from scratch since it is too error-prone. Libraries with single building blocks and ready-to-use mechanisms are the preferred alternative. There are several libraries available that provide the basic mechanisms for different programming languages (e.g. Python, Java, C, GO, C++). We give a brief overview about some of them.

A library that enables the application of basic mechanisms is the *Google Differential Privacy library* [37]. With this library at the same time, the mechanisms can be built on existing frameworks such as Apache Beam. The *Google RAPPOR library* [29] enables the application of the RAPPOR algorithm which was previously presented in the context of the Randomized Response Method. The *OpenDP* project [38] also provides easy access to the mechanisms to apply them to individual data with the *smartnoise library*. *IBM* offers a library to use the discrete Gaussian mechanism [39]. *Ektelo* is also a framework for implementing privacy algorithms [40].

Furthermore, libraries exist that allow the training of neural networks under conditions of differential privacy. The *Opacus* [41] framework allows differentially private training of PyTorch models. *Opacus* is open source and offers a modular API. *TensorFlow* [42] also offers a library that enables the differential private training of neural networks. The *OpenMined Project* [43], with its *Syft* and *Grid* modules, applies differential privacy in the context of federated learning. The project is compatible with PyTorch and TensorFlow. The *diffprivlib* [44] by *IBM* offers basic mechanisms, which can be applied individually by the user. However, simple machine learning algorithms such as a random forest or a logistic regression can also be trained under differential privacy with the *diffprivlib*.

Since it is not trivial to determine how the algorithms are implemented in detail in each library, a comparison was conducted by Garrdio et.al [45]. The libraries were compared qualitatively, and quantitatively with four different types of queries using synthetic and real-world datasets. All libraries were suitable for productional use, however, they differ strongly in the function range. However, no library satisfies a universal utility for all applications. [45]

6. Directions for future research

The practical examples in section 5 show that specific mechanisms are needed considering different variable types. Extending the mechanisms to publishing multidimensional data, i.e. mixed data containing numeric and categorical data types, is not trivial. Research shows that applying the mechanisms to the individual attributes yields poor results. Therefore, solutions must be developed that can perturb multidimensional datasets in total containing numeric and categorical variables with the optimal worst-case error. [46–48].

The given examples also show that different queriers who act independently externally but combine their knowledge gained later, must be taken into account during the design of a LDP system. This comes into effect if other third parties can have access to the data instead of just the machine manufacturer. In case of doubt, the number of requests from each querier and the number of same requests must be limited [49]. If each analyst receives a slightly different answer, analysts could collaborate and calculate the mean of their answers. In a worst-case scenario, they are able to determine the true value. In this case, it makes sense to limit the number of queries and send the same answer to each analyst [28]. After defining the collaborators and the variable types to be published have been determined, the mechanisms for adding the noise must be suitably parameterized. Since the parameter epsilon ε is a measure of privacy and is also needed for the parameterization of the mechanisms, the choice of this value is very crucial. Currently there is no best practice for setting ε for a desired privacy utility tradeoff. Thus it would be helpful if early adopters of differential privacy could share their ε values from real-world applications [50].

In the long term, it would be desirable to be able to publish entire differentially private datasets [49]. In the context of Open Science, the release of whole datasets would also be advantageous. Companies could share their data with the machine learning community without having any privacy concerns. The use of machine learning and artificial intelligence would become more quickly applicable through the collaborative work of the community and thus take a further step toward autonomous production.

7. Conclusion

By identifying and defining an application scenario, mapping the concepts of LDP to the manufacturing context, we have shown that the demands and potential threats to privacy leaks when publishing or sharing data with third parties are of a different kind compared to the threats when considering public datasets for the protection of individual personal data. It must be understood that the company's process knowledge can be leaked by sharing production data with third parties. Beyond important process information, which is necessary to produce cost-effective products, sensitive business data as well as strategic data can be revealed. To apply LDP in the manufacturing context, it is mandatory to analyze the use case in advance. It should be asked who will have access to what kind of data and which potential threats can arise by sharing the data. From the point of view of the curator (machine manufacturer), it must be taken into account that the data amounts must be correspondingly large in order to learn valid insights. In general, when publishing differentially private data, it must be taken into account that there is a tradeoff between accuracy and privacy. There is no generic approach for determining the ideal value of the parameter yet.

Acknowledgments

This work was partially supported by a grant from the Federal Ministry of Education and Research providing for this work through the project “DigiBattPro 4.0” and from the State Ministry of Baden-Wuerttemberg for Economic Affairs, Labour and Tourism (Project *“KI-Fortschrittszentrum: Lernende Systeme und Kognitive Robotik”*).

References

[1] Arinez, J.F., Chang, Q., Gao, R.X., Xu, C., Zhang, J., 2020. Artificial Intelligence in Advanced Manufacturing: Current Status and Future Outlook. Journal of Manufacturing Science and Engineering 142 (11).

[2] Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. CIRP Annals 65 (2), 621–641.

[3] Wang, L., 2019. From Intelligence Science to Intelligent Manufacturing. Engineering 5 (4), 615–618.

[4] Savazzi, S., Nicoli, M., Bennis, M., Kianoush, S., Barbieri, L., 2021. Opportunities of Federated Learning in Connected, Cooperative, and Automated Industrial Systems. IEEE Commun. Mag. 59 (2), 16–21.

[5] Arvind Narayanan, Vitaly Shmatikov, 2006. Robust De-anonymization of Large Datasets (How to Break Anonymity of the Netflix Prize Dataset).

[6] M. Douriez, H. Doraiswamy, J. Freire, C. T. Silva, 2016. Anonymizing NYC Taxi Data: Does It Matter?, in: 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA). 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp. 140–148.

[7] Pierangela Samarati, L.S., 1998. Protecting Privacy when Disclosing Information: k-Anonymity and Its Enforcement through Generalization and Suppression.

[8] Liu, K., Giannella, C., Kargupta, H., 2008. A Survey of Attack Techniques on Privacy-Preserving Data Perturbation Methods, in: Aggarwal, C.C., Yu, P.S. (Eds.), Privacy-preserving data mining. Models and algorithms, vol. 34. Springer, New York, NY, pp. 359–381.

[9] Singh, N., Singh, A.K., 2018. Data Privacy Protection Mechanisms in Cloud. Data Sci. Eng. 3 (1), 24–39.

[10] Martin M. Merener, 2012. Theoretical Results on De-Anonymization via Linkage Attacks. Transactions on Data Privacy 5 (2), 377–402.

[11] Naveed, M., Kamara, S., Wright, C.V., 2015. Inference Attacks on Property-Preserving Encrypted Databases, in: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. CCS'15: The 22nd ACM Conference on Computer and Communications Security, Denver Colorado USA. 12 10 2015 16 10 2015. ACM, New York, NY, pp. 644–655.

[12] Rigaki, M., Garcia, S., 2020. A Survey of Privacy Attacks in Machine Learning. https://arxiv.org/pdf/2007.07646.

[13] Dwork, C., 2006. Differential Privacy, in: Hutchison, D., Kanade, T., et.al. (Eds.), Automata, languages and programming. 33rd international colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006; proceedings, vol. 4052. Springer, Berlin, pp. 1–12.

[14] A. G. Thakurta, A. H. Vyrros, U. S. Vaishampayan, G. Kapoor, J. Freudiger, V. R. Sridhar, and D. Davidson. Learning new Words.

[15] Differential Privacy Team, 2017. Learning With Privacy at Scale.

[16] Messing, S., DeGregorio, C., Hillenbrand, B., King, G., Mahanti, S., Mukerjee, Z., Nayak, C., Persily, N., State, B., Wilkins, A., 2020. Facebook Privacy-Protected Full URLs Data Set.

[17] Braud, A., Fromentoux, G., Radier, B., Le Grand, O., 2021. The Road to European Digital Sovereignty with Gaia-X and IDSA. IEEE Network 35 (2), 4–5.

[18] Siderska, J., Jadaan, K.S., 2018. Cloud manufacturing: a service-oriented manufacturing paradigm. A review paper. Engineering Management in Production and Services 10 (1), 22–31.

[19] Wang, K., 2006. Data Mining in Manufacturing: The Nature and Implications, in: Wang, F., Wang, K., Kovacs, G., Wozny, M., Fang, M. (Eds.), Knowledge enterprise: intelligent strategies in product design, manufacturing, and management. Proceedings of PROLAMAT 2006, IFIP TC5 international conference, June 15-17 2006, Shanghai, China, vol. 207, 1. Ed. ed. Springer, New York, NY, pp. 1–10.

[20] Mahawaga Arachchige, P.C., Bertok, P., Khalil, I., Liu, D., Camtepe, S., Atiquzzaman, M., 2019. Local Differential Privacy for Deep Learning 7. https://arxiv.org/pdf/1908.02997.

[21] D Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Dan Dennison, 2015. Hidden Technical Debt in Machine Learning Systems. Advances in Neural Information Processing Systems, 2494–2502.

[22] Wang, K., 2007. Applying data mining to manufacturing: the nature and implications. J Intell Manuf 18 (4), 487–495.

[23] Aggarwal, C.C., Yu, P.S. (Eds.), 2008. Privacy-preserving data mining: Models and algorithms. Springer, New York, NY, 513 pp.

[24] Hassan, M.U., Rehmani, M.H., Chen, J., 2020. Differential Privacy Techniques for Cyber Physical Systems: A Survey. IEEE Commun. Surv. Tutorials 22 (1), 746–789.

[25] Yang, M., Lyu, L., Zhao, J., Zhu, T., Lam, K.-Y., 2020. Local Differential Privacy and Its Applications: A Comprehensive Survey, 24 pp. https://arxiv.org/pdf/2008.03686.

[26] Warner, S.L., 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. Journal of the American Statistical Association 60 (309), 63.

[27] Erlingsson, Ú., Pihur, V., Korolova, A., 2014. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response, 14 pp. https://arxiv.org/pdf/1407.6981.

[28] Leoni, D., 2012. Non-interactive differential privacy, in: Proceedings of the First International Workshop on Open Data. the First International Workshop, Nantes, France. 5/25/2012 - 5/25/2012. ACM, New York, NY, p. 40.

[29] Google. RAPPOR. https://github.com/google/rappor. Accessed 3 February 2022.

[30] Bassily, R., Smith, A., 2015. Local, Private, Efficient Protocols for Succinct Histograms, in: Proceedings of the forty-seventh annual ACM symposium on Theory of computing. STOC '15: Symposium on Theory of Computing, Portland Oregon USA. 14 06 2015 17 06 2015. ACM, New York, NY, pp. 127–135.

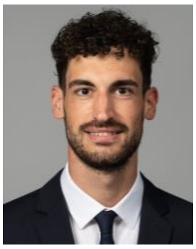
[31] Tianhao Wang, Jeremiah Blocki, and Ninghui Li, Somesh Jha. Locally Differentially Private Protocols for Frequency Estimation, in: , Proceedings of the 26th USENIX Security Symposium, vol. 26.

[32] S. Xiong, A. D. Sarwate, N. B. Mandayam, 2016. Randomized requantization with local differential privacy, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2189–2193.

[33] Ye, Q., Hu, H., Li, N., Meng, X., Zheng, H., Yan, H., 2021. Beyond Value Perturbation: Local Differential Privacy in the Temporal Setting. INFOCOM 2021 - IEEE Conference on Computer Communications, 1–10.

[34] Bi, M., Wang, Y., Cai, Z., Tong, X., 2020. A privacy-preserving mechanism based on local differential privacy in edge computing. China Commun. 17 (9), 50–65.
[35] Jiang, K., Shao, D., Bressan, S., Kister, T., Tan, K.-L., 2013. Publishing trajectories with differential privacy guarantees, in: Proceedings of the 25th International Conference on Scientific and Statistical Database Management. the 25th International Conference, Baltimore, Maryland. 7/29/2013 - 7/31/2013. ACM, New York, NY, p. 1.
[36] Kim, J.W., Kim, D.-H., Jang, B., 2018. Application of Local Differential Privacy to Collection of Indoor Positioning Data. IEEE Access 6, 4276–4286.
[37] Google. Google Differential Privacy. https://github.com/google/differential-privacy. Accessed 3 February 2022.
[38] OpenDP. smartnoise-sdk. https://github.com/opendp/smartnoise-sdk. Accessed 3 February 2022.
[39] IBM. discrete gaussian differential privacy. https://github.com/IBM/discrete-gaussian-differential-privacy. Accessed 3 February 2022.
[40] Ektelo. Ektelo. https://github.com/ektelo/ektelo. Accessed 3 February 2022.
[41] Opacus. Opacus. https://opacus.ai/. Accessed 3 February 2022.
[42] TensorFlow. TensorFlow Privacy. https://github.com/tensorflow/privacy. Accessed 3 February 2022.
[43] OpenMined. PySyft. https://github.com/OpenMined/PySyft. Accessed 3 February 2022.
[44] IBM. differential-privacy-library. https://github.com/IBM/differential-privacy-library. Accessed 3 February 2022.
[45] Garrido, G.M., Near, J., Muhammad, A., He, W., Matzutt, R., Matthes, F., 2021. Do I Get the Privacy I Need? Benchmarking Utility in Differential Privacy Libraries, 13 pp. https://arxiv.org/pdf/2109.10789.
[46] Nguyên, T.T., Xiao, X., Yang, Y., Hui, S.C., Shin, H., Shin, J., 2016. Collecting and Analyzing Data from Smart Device Users with Local Differential Privacy, 11 pp. https://arxiv.org/pdf/1606.05053.
[47] Wang, N., Xiao, X., Yang, Y., Zhao, J., Hui, S.C., Shin, H., Shin, J., Yu, G., 2019. Collecting and Analyzing Multidimensional Data with Local Differential Privacy. https://arxiv.org/pdf/1907.00782.
[48] Wang, T., Ding, B., Zhou, J., Hong, C., Huang, Z., Li, N., Jha, S., 2019. Answering Multi-Dimensional Analytical Queries under Local Differential Privacy, in: Proceedings of the 2019 International Conference on Management of Data. SIGMOD/PODS '19: International Conference on Management of Data, Amsterdam Netherlands. 30 06 2019 05 07 2019. Association for Computing Machinery, New York,NY,United States, pp. 159–176.
[49] Mohammed, N., Chen, R., Fung, B.C., Yu, P.S., 2011. Differentially private data release for data mining, in: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. the 17th ACM SIGKDD international conference, San Diego, California, USA. 8/21/2011 - 8/24/2011. ACM, New York, NY, p. 493.
[50] Dwork, C., Kohli, N., Mulligan, D., 2019. Differential Privacy in Practice: Expose your Epsilons! JPC 9 (2).

Biography



Sascha Gärtner (*1993), is a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation (IPA) in Stuttgart, Germany. As part of his work in the Competence Center for Digital Tools in Production, he researches the practical application of artificial intelligence methods in the smart manufacturing environment.

Michael Oberle (*1984) is a computer scientist and group leader working in the field of smart manufacturing with a focus of data-driven and event- driven production control services. One of his main achievements is the fully connected cloud-controlled battery manufacturing pilot line at Fraunhofer IPAs technical centre for battery manufacturing.

3rd Conference on Production Systems and Logistics

Generation of a Data Model For Quotation Costing Of Make To Order Manufacturers From Case Studies

Katharina Berwing[1], Günther Schuh[1], Volker Stich[1]
[1] FIR at RWTH Aachen, Aachen, Germany

Abstract

For contract or make to order manufacturers, quotation costing is a complex process that is mainly performed based on experience. Due to the high diversity of the product range of these mostly small or medium-sized companies (SMEs) and the poor data situation at the time of quotation preparation, the quality of the calculation is subject to strong variations and uncertainties. The gap between the initial quotation costing and the actual costs to be spent (pre- and post-calculation) is crucial to the existence of SMEs. Digitalization in general can help companies to get a better understanding of processes and to generate data. For improving these processes, an understanding of the important data for that specific process is crucial. Accurate quotation costing for customized products is time-consuming and resource-intensive, as there is a lack of an overview of data to be used within the process. This paper therefore derives a data model for supporting quotation costing in the company, based on literature-based costing procedures and recorded case studies for quotation and calculation. Based on the results, SMEs will have a first overview of the needed data for quotation costing to optimize their calculation process.

Keywords

Calculation costing; pre- and post-calculation; data model; make to order manufacturer

1. Introduction

As stated above, the quotation costing is a very manual and time-consuming process with often inaccurate results for make to order manufacturers. The general digitalization trend can help these companies to improve the results by using valid and extensive data. As a first step to do so, companies must know, which data are important for the calculation process. The data can be shown in a data model derived from case studies of different make to order manufacturers to learn, how equal or different they use data for the quotation costing process.

Today, companies in all industries are faced with the major challenge of digitizing their processes and using IT systems so that business processes can be carried out more automatically and efficiently. IT is assuming the role as an enabler of digitization in this process. [1] With the trend of digitalization in companies, the value of data has also been recognized. [2] The availability of information, and thus the ability to generate knowledge from data, significantly determines the competitiveness of a company. [3] IT systems in use require and generate data, which in turn can be used to optimize business processes and act more economically. However, there are some processes that still require a lot of manual work and personal experience. One of these is the process for quotation calculation in companies that produce individual products for customers (make to order manufacturers). To improve that complex and time-consuming process, an understanding of the actual data used and a generation of a data model with all needed data is

DOI: https://doi.org/10.15488/12160 publish-Ing.

* The paper successfully passed the peer review process but was not presented at a conference.

the first step to reach a data-based calculation. This paper deals with identifying the used data for quotation calculation from case studies and deriving a data model to optimize that process.

2. Theoretical background

In the following, the general problem statement and the theoretical foundations are described as well as the state of the art in the field of quotation costing for SMEs.

2.1 Problem statement and theoretical foundations

Calculation costing determines the costs for the production of a product or service unit. While direct costs can be allocated to an individual product or order, the allocation of overhead costs according to their cause is a major challenge for contract manufacturers. [4] [5] Quotation costing is the calculation for the first quotation for the customer. The challenge in quotation costing for contract manufacturers is that the quotation should be customer-oriented on the one hand, by determining the customer's upper price limits, and cost-oriented on the other hand, by determining the cost-covering lower price limit. [6] Companies with a high degree of complexity, such as contract manufacturers, require a much more detailed cost accounting system in order to be able to compete in the market than companies with a lower degree of complexity. However, a study shows that this is not implemented in practice, resulting in mostly inaccurate cost statements. [7]

Make to order manufacturers use different methods for quotation costing (kilo-cost method, material cost method, division costing, similarity costing, equivalence number costing, surcharge costing, target pricing), which, however, are all characterized by a lack of precision due to the large proportion of estimated and empirical values. [8] A study on quotation costing shows that overhead costing is by far the most frequently used costing method. [7] However, this highly simplified method only allows a limited allocation of cost by its causes. The specific procedure for identifying the prime cost for the quotation calculation differs in the practical application in different companies, the theoretical goal however is always to find the exact costs that are spent to produce a good.

A fast and precise quotation calculation represents an important competitive factor in industrial sectors, due to the fierce competition, the result must already have a very high quality. [9] To be able to allocate costs according to causes would therefore mean an enormous gain regarding the quality of the quotation costing. Activity-based costing aims to allocate costs according to their cause and can thus reveal inefficiencies. But this type of cost accounting is only suitable for repetitive tasks and thus cannot be implemented in a meaningful way as a costing methodology for make to order manufacturers. [10] With increasing product variety and differences in the number of units, the allocation of costs within the company according to their cause becomes increasingly complex and difficult to implement economically. [11] For the calculation of a quotation this means that in contradiction to the great uncertainty of cost statements at the beginning of a project, there is the requirement of exact cost target setting. [11]

2.2 State of the Art

With the help of the repeatability of calculation results, it is possible to assess the calculation results in terms of their quality. The degree of standardization of methods correlates positively with the repeatability. A study shows that while 90% of respondents attach high or very high importance to repeatability, less than a third of companies are able to achieve this goal. Consequently, repeatability and the degree of standardization are considered significant by companies, but they are only partially able to meet this goal. [7]

Based on an automatic analysis of CAD models of a work piece without detailed work and production planning, KNOBLACH has developed a system for a fast quotation calculation for flexible processes of sheet metal part production and for production with progressive dies. [12] This procedure finds an application exclusively for sheet metal parts with specific process and manufacturing peculiarities (consideration of procedures of the shearing and laser cutting). All in all, the approach becomes very company-specific, since the respective company-specific cost structure must be represented in the calculation scheme. Another possibility is the approach of HAUSCHILD, who dealt with dynamic quotation costing for contract manufacturers. The aim of his work is to improve the internal coordination of the departments involved in costing (production, sales, purchasing). In addition, he has taken into account the time variability of the input variables used in his approach in order to increase the meaningfulness of the quotation costing. [13] In this approach a strong focus is put on the temporal factors of the in- and out payments and the data needed and available for a quotation costing is rather neglected. Due to the only sporadic use of corporate IT at the time of both publications, the data available in times of a variety of production systems are not taken into account, which leads overall to a non-transferable approach.

With the help of special costing procedures, the costs for geometrically similar parts can be identified by providing geometrically based cost forecasts. [14] [15] In the application, an IT-supported, database-based costing model for project-accompanying costing has been developed, which consists of a coupling of methods for processing technical and business information and product costing. The procedure is based on feature descriptions of the production parts and uses the procedure of Case Based Reasoning. [16] However, the approach is not suitable for evaluating different components or for comparing different manufacturing processes. Furthermore, the approach focuses on product-accompanying costing and not on the process of quoting before manufacturing starts.

In addition, there are also approaches that use artificial intelligence algorithms to approach the topic of manufacturing costs. However, these refer to other areas of application that can be calculated with simpler rules than make to order manufacturing (e.g. additive manufacturing, mass customization). [17] [18]

WESTEKEMPER has looked into the methodology of quotation pricing for contract manufacturers. Initial costing procedures can also be found here, but the research focus is on pricing procedures, which addresses a completely different focus. The costing procedures he stated do not go beyond the state of the art. [19]

A final look at the existing software landscape in the area of quotation costing shows that the existing software supporting this process either focuses on a specific industry (mainly automotive), or does not contain its own data model for the relevant data to be used, but transfers the previously manually calculated criteria into a software environment. This, in turn, is to be seen more as a digitization step and does not improve the general process.

In conclusion, it is clear from the state of the art that there are a variety of approaches for supporting quotation calculation for contract manufacturers. What is lacking so far, is a data-driven approach that supports a direct improvement of the general calculation process as well as the calculation result through an easy application.

3. Methodological approach

In order to develop a data-supported or data-based quotation calculation, the data and criteria used by different companies within the quotation process are recorded in the methodical procedure by means of case studies. These different data are merged in a data model, which enables a later application in an IT system and thus represents a first step towards an actual data-based quotation costing. The research question can be formulated as follows:

"What data is used in the quotation costing process of different make to order manufacturers and how does a data model look like based on these information?"

Research Procedure

First, case studies have been conducted in semi-structured interviews. The companies for the case studies were selected according to the following criteria: customer-specific production, complex product structure, make to order manufacturer (one-time or small batch production), small or medium-sized companies. In addition, a focus was placed on the diversity of the different companies (degree of digitization, company size, products, etc.) in order to obtain the most comprehensive result possible for the data model. During the interviews, after a brief theoretical introduction, the quotation costing process of the specific company was presented and discussed in detail. In each case, particular focus was placed on the data and associated IT systems used in individual process steps. At the end of the interview, the main challenges for the company in quotation costing and the information needed for optimizing the process were discussed. As a result, the case studies revealed the data used for quotation costing with the respective source (personal knowledge, Excel, ERP system, etc.) and further information and data needed for improving the quality of the quotation costing result. Based on these results, a data model was derived for each case study. Subsequently, these data models have been extended and combined into a common data model for quotation costing.

Background Data Model

A data model is a *"model of the data to be described and processed in an application area and their relationships to each other"*. [20] Data models have emerged from the desire to organize existing data. As the use of IT systems increases, so does the volume of existing data and the demands on the quality of this data. With the help of ordering criteria and distinguishing features, data models structure relevant data. [21] With the help of this approach, the transparency of organizational structures is increased and potential for improving processes in the company is revealed, making data models an aid for solving organizational issues. [3] The logical data model is a data model for representing and explaining the statistical and database elements of a business unit or the requirements for its procedures and techniques in a logical and theoretical way that eventually leads to their application in a database. [22]

Unified modelling language (UML) is a standardized representation or notation that allows object-oriented models to be represented in a uniform manner. Class diagrams can be used to represent classes and objects, their attributes and methods, and the relationships (so-called associations) between them. Classes and objects are represented as boxes. As shown in Figure 1, a class always contains a class name at the top. In the second section of the class, in the middle area, attributes are listed that can be assigned to the class. Behind a ":" likewise the data type (string, int, boolean...) is indicated. [23] Beneath this the methods are listed. Objects, which are specific examples of a class, possess a name, which is underlined in order to be able to differentiate an object quickly and transparently from a class. [24]

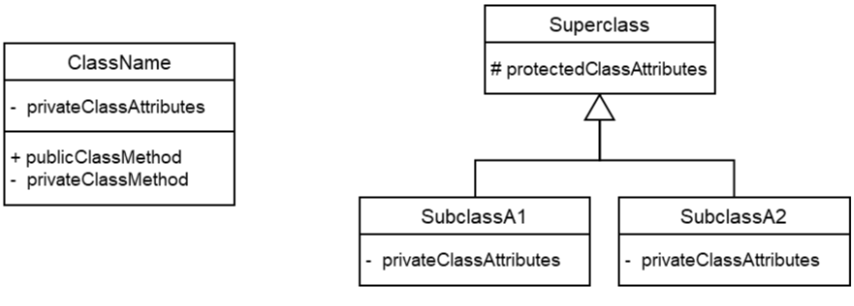


Figure 1: Example of UML Class Diagram

The visibility of attributes or methods of a class can also be mapped in UML. If an element of a class is public, this means that the corresponding element is also visible outside the class. Accordingly, if the attribute or method is prefixed with a "+". A "-" is used if the element should only be visible within its own class (private). Finally, the property protected can be listed, which describes that an element is only visible within the class in which it is described and in all subclasses. [24] The relationships that exist between the classes and/or objects are represented with the help of lines. Here, three different types of relationships are distinguished, associations, whole/part relationships and general/specialization relationships. [24]

4. Case studies

Different companies were interviewed in order to identify relevant and used data from SMEs for quotation costing. These companies have in common, that they all are contract manufacturers, which means that they face challenges for complex products with small batch sizes. On the other hand, these companies can be classified as small or medium-sized enterprises for which the quality of the quotation costing plays an important role. In order to also reflect the diversity within this group of companies, the companies differ in the type and scope of their service provision (products, services, interaction with customers, degree of digitization, etc.). The aim of the interviews were to get an understanding of the process of cost calculation in the context of quotation costing. In particular, it was asked, which criteria are considered to be cost-relevant variables and which decision-making criteria are used to determine the level of costs. First and foremost, experience is a highly relevant aspect when estimating costs at an early stage, even before the customer actually places an order. Therefore, an attempt was made to list criteria that are relevant for decision-making when using experience-based, manual adjustments.

Company 1 – Interview conducted with the General Manager

The company was founded in 1969. It is a DIN ISO 9001 certified company that manufactures prototypes, individual parts and small series. The company specializes in particular in the machining of components, some of which are large. In addition to machining, they also offer their own design department (machine and plant construction) as well as steel construction work including welding technology.

The costing process at Company 1 begins with a customer inquiry, which in turn triggers an internal feasibility analysis. The product specifications (e.g. component dimensions, necessary production processes and number of pieces) are used. On the one hand, this enables a check to be made as to whether the technological constraints for manufacturing the components can be met within the company and whether all the relevant information for a complete production order is available. It is also checked on the basis of the current production capacity utilization, which is derived from the Enterprise Resource Planning System (ERP system), whether sufficient capacities are available in production to accept the order. After a successful check, a corresponding production order is created, whereby a work schedule and the associated bill of materials are generated. This information is now used to calculate the cost of goods sold as part of a preliminary costing. The cost of goods sold is made up of four components. These include the machine hourly rate calculation, the calculation of material costs, the consideration of external costs and other costs.

First, the production costs are calculated with the help of a spreadsheet calculation in Excel. Here, the work schedule is used and the planned machine utilization is offset against the corresponding machine hourly rates. The machine hourly rates are based on the usual variables of the hourly rate calculation (investment costs, depreciation, maintenance costs, etc.), but in this case also on an overhead surcharge and the production wages. As a result, this cost module can be used to calculate all costs incurred during production on the corresponding machines.

For the calculation of the material costs, the bill of material is used. The required raw materials are selected on the basis of dimensioning, material selection and quality and the corresponding quantity is offset against the current material prices. This then results in the material costs. It was emphasized at this point that the increasing volatility of material prices plays a significant role in this cost module. As a rule, shipping costs are not incurred, as customers often prefer to collect the goods themselves. However, if the customer requests delivery, a shipping surcharge is calculated based on the transport volume and the distance to the customer. The level of these costs is also influenced by whether long-term customer relationships already exist.

After adding up the cost areas described above, the cost price for the requested component is calculated. A profit markup and any cash discount are added to these cost prices to obtain the net quotation price. In addition to this, the earliest possible delivery date is also specified and sent to the customer within a quotation. This completes the quotation process.

Mainly used criteria for quotation costing: Product specifications, production capacity check, work schedule, bill of material, machine hourly rate, material costs out of current material prices

Company 2 – Interview conducted with the Production Controller

As a subsidiary of a larger technology group Company 2 is a producer of banknotes and security papers. The company was founded in 1964 and today supplies banknotes to the European Central Bank.

The request to the company includes the product specifications, which include the desired paper/substrate type and the number of pieces. Once all the relevant information are available, a production order is created, which includes the work schedule and bill of materials. The quotation calculation then follows using a spreadsheet containing current market prices for the required material, the annually updated planned values, and cost curves based on post-calculations of sold products. The product specifications are entered into the calculation tool and the cost curve provides information on the expected production costs depending on the batch size. The material costs are also derived accordingly from the spreadsheet. A profit markup is added in the end. Numbers may be corrected again manually. This is done on the basis of findings from variance analyses, which are carried out monthly. An attempt is made to identify and understand deviations in past customer orders.

Mainly used criteria for quotation costing: Product specifications, work schedule, bill of material, current market prices, cost curve from post-calculation of sold products

Company 3 – Interview conducted with the Sales Representative

Company 3 was founded in Cologne in 1896 and offers a wide range of printed products for commercial purposes. In addition to the product steps of image processing and advertising technology, the company offers in particular digital and screen printing as well as offset printing.

The process of quotation calculation at Company 3 starts with the customer inquiry, which includes product specifications such as the product type (e.g. advertising banner), the motif and the number of pieces. If the necessary information is complete, a production order with a corresponding work schedule is derived from it. The calculation consists of the direct production costs, the material costs and the additional overhead costs. The work schedule and the resulting machine assignment serve as the basis for the direct production costs. The machine hourly rate associated with the respective machines is then multiplied by the corresponding forecast production time per machine to calculate the direct production costs. Here, production times are based to a large extent on the experience knowledge of the calculator. Important factors are the printing speed, which is strongly dependent on the selected motif, and estimates of the production employee. The material unit costs result from the production order or the customer's inquiry. A material selection is made, whereby the material quantity in turn depends on the work schedule. This is because, depending on the machine selection and the associated offcuts, as well as the material scrap during setup of the machine, there is an additional requirement for raw material. Consequently, the total material requirement is determined by the calculator and offset against the material prices. The material prices are again stored in the system and are based on price agreements requested on a quarterly basis. Finally, the total direct material costs are calculated.

Together with a fixed surcharge for administration and sales, which is independent of the order volume, this results in the quotation cost. A profit markup, a cash discount and any commissions are then added to these to obtain the net quotation price.

Mainly used criteria for quotation costing: Product specifications, work schedule, material cost, overhead cost, machine hourly rate, machines used, forecast of production time, material prices for production material

Company 4 – Interview conducted with the Operations Planning Representative

Company 4 was founded in 1996 and specializes in the manufacturing and repair of rolls. In addition to manufacturing, disassembly and repair, the company also offers surface treatment and coating of rolls and specialized technical consulting.

After the process starts with a customer inquiry, the product specifications of the inquiry are analyzed. These include in particular the technical drawing of the product to be manufactured, where the component dimensions and the required manufacturing processes can be derived from. A corresponding production order is created, from which a work schedule can be taken. The quotation calculation then follows in the sense of a differentiated markup calculation. A distinction is made between the cost items direct production costs, direct material costs, special direct sales costs, administrative overheads and production overheads. Comparable orders are considered in the calculation and individual numerical values are estimated. The actual work schedule of the products already manufactured are exported from the accounting system.

First, the direct production costs are determined. For this, the machine hourly rates and production wages are taken as a basis and multiplied by the respective forecast production times. The work schedule is used as a basis to map the machine selection. The manual input of the production times is of great importance in order to ensure a high degree of accuracy in the cost calculation. The times of comparable orders stored in the booking system are initially used as the basis for estimating the production time. Based on this, empirical values are used to increase precision. The calculator takes into account which employee is listed as a machine operator in the booking system. For example, the experience of long-serving employees can be an indicator that the production job already completed was completed particularly quickly. A more conservative estimation of the production time by adding a corresponding markup could therefore be recommended. Also the posting date in connection with the considered machine tool is used as criterion for the estimation of the production times. If a machine is newly acquired, it can be expected during commissioning that the employee is still untrained in its operation. Delays caused by technical problems that the machine operator is not able to eliminate as quickly as expected in normal operation are also conceivable. The transferability of the production times could therefore be accompanied by a certain lack of precision when this information is taken into account. A manual correction is therefore necessary.

Another criterion is parallel jobs that may have been performed on the same machine without this being recorded in the booking system by the employee. This circumstance is not directly apparent, but can be checked by looking at other jobs that were logged in the booking system on the same day.

In addition, the work schedule of the historical data under consideration can provide information about problems that the machine operator had to rectify. If incidents were documented, the calculator can make an appropriate correction to the production times on this basis. Material errors or setup times are conceivable incidents that delay further processing and can thus greatly distort the production time booked in the system. Once the direct production costs have been calculated by multiplying the respective hourly machine rates (or the equivalent of production wages) by the forecast production times, the direct material costs are determined. The necessary raw materials can be taken from the bill of materials of the production order. However, the design drawing also provides information on the selection of raw materials. The tolerances required by the customer and the mechanical boundary conditions are used and compared with the raw materials or standard parts actually available. A correction of the parts list is often necessary. Multiplying the material selection by the corresponding material prices finally yields the material unit costs. The material prices are determined on the basis of quarterly price agreements. In this calculation, it should be noted that purchased parts are not yet taken into account. These are declared as special direct costs of the sales department and result either from concrete offers from suppliers, are estimated as a lump-sum markup or can also be taken from post-calculations of past orders. The calculated costs are then added together and a variable administrative and production overhead rate is applied. The result is the cost of goods sold of the costing object. After taking into account the profit markup and a cash discount, the result is the net quoted price.

Mainly used criteria for quotation costing: Product specifications (technical drawing of product), work schedule, comparable orders, machine hourly rate, production wages, forecasted production time (from work schedule of historical comparable orders), machine operator for historical orders, experience of machine operators, machine tool, check of parallel jobs of historical common orders, incidents documented in the work schedule

5. Data model for quotation costing

For the generation of the data model for quotation costing, the results from the interviews are summarized and explained briefly in the following.

The expert interviews show that the criteria used for the quotation and the cost calculation is highly agreed upon by the interviewees. A production order is first created on the basis of the customer inquiry, which consists of the bill of material and the work schedule. This is then followed by the calculation of the cost. Here, a rough distinction can be made between production costs, material costs and other overheads.

The production costs consist of those costs which arise from the machine assignment in the production. The basis for this is the work schedule, it specifies the sequence and allocation in which the raw parts are produced on the respective machines and tools. This can be used for a machine hourly rate calculation, the cost rates for the machines are fixed. A decisive variable for the cost determination is the production time. This is determined by the calculator as stated in the case studies on the basis of various factors. A correct estimation of the production time is a decisive success criterion for the precision of the forecast cost calculation. The interviews show that in business practice, empirical values are of great importance. Production times are estimated in conjunction with discussions with production employees and on the basis of (manually filtered) historical data. These are then multiplied by the corresponding machine hourly rates to produce the production costs.

A preliminary calculation of the expected material costs is based merely on the evaluation of the bill of materials. It provides information on which raw parts and purchased parts are required for the production order. In addition to the dimensioning, type and number of pieces, it also shows which material specifications exist in order to be able to make an appropriate material selection. Depending on the industry and product spectrum, corrections may have to be made here because the requested components do not correspond to the dimensions of unfinished parts. This process, which is also based on personal experience, takes place with the help of feedback from the customer. Once the bill of material is finalized, costs are calculated and totaled based on price agreements, lump sums, or specific quotes from suppliers to determine the total material costs. The production costs are added to the material costs to give the expected production costs. In order to include those costs in the quotation that are not directly attributed to the criteria stated above, an overhead markup is added to the production costs. This is usually a fixed cost rate that is calculated based on an overhead calculation. In particular, administrative and selling expenses are included. All these information are summarized and outlined in the following UML Diagram (Figure 2).

Description of the UML Diagram

The class model “Prime Costs” is used to calculate the cost of goods and the UML notation makes it possible to quickly obtain a general understanding of the basic structure of the data model. The prime cost class is at the end of the calculation. This class is used to calculate the quotation cost for the customer on the basis of direct and overhead costs. The latter are determined on the basis of a stored overhead rate on the direct costs. The direct costs in turn consist of the material and production costs.

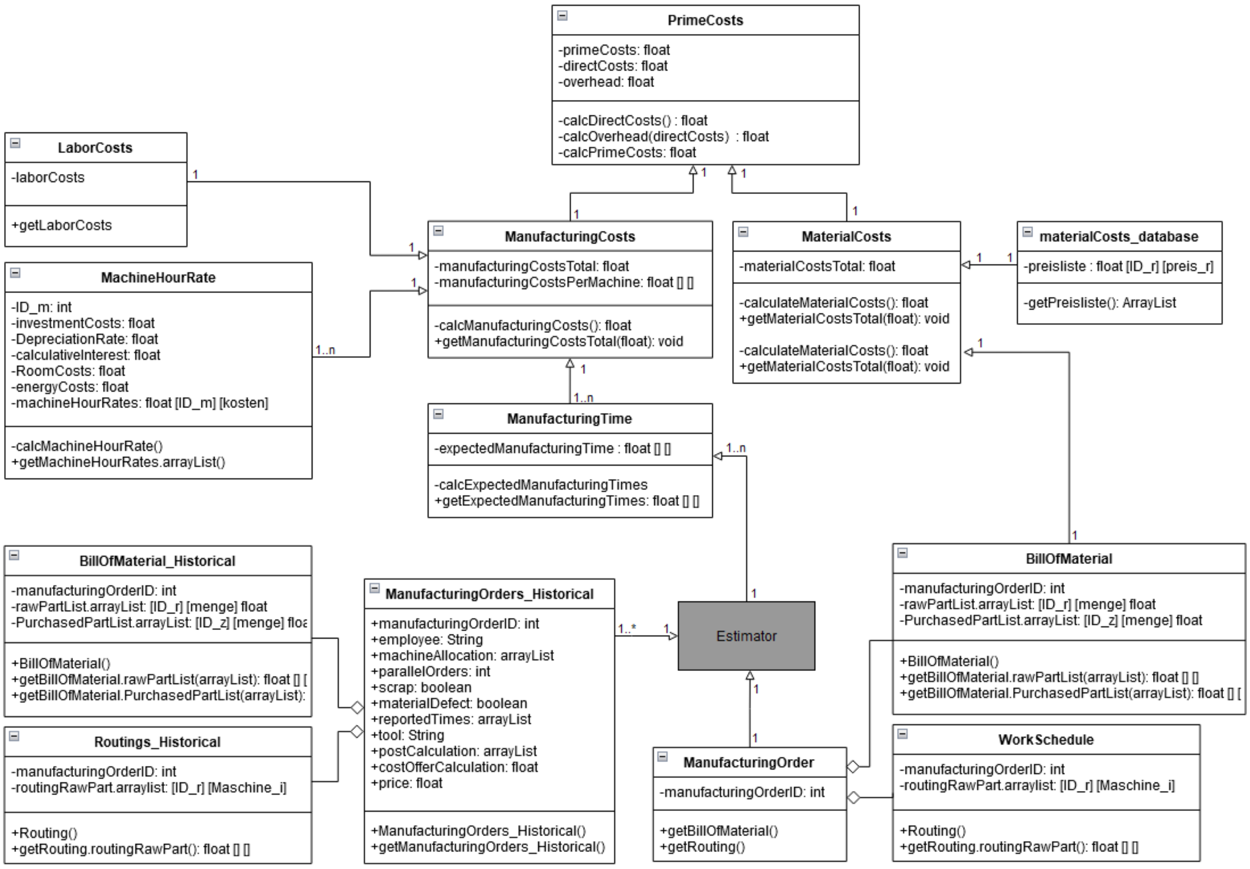


Figure 2: Data Model for Quotation Costing in SMEs

In the class material costs, the information from the bill of materials (BoM) is related to the costs for the corresponding raw parts and purchased parts. After creating an array list that associates the items from the BoM with the corresponding material prices from the Material Database class, the total material costs can be returned using a simple multiplication. This value is stored in the materialCostsTotal attribute and can be retrieved by a corresponding getter method in the cost price class.

The production costs are calculated in the corresponding class with the help of machine hourly rates, the wage costs and the predicted production times. The machine hourly rates are also initialized as a separate class. This has all the relevant attributes of each machine that is actually used within the production operation. In addition to the machine-specific ID, the class stores all relevant values for the complete calculation of the machine hour rates. The result is stored in an array list containing the machine ID and the machine costs for one hour for all machines. The labor costs in turn are set as attribute laborCosts. A corresponding method in the production costs class can retrieve the values. The production time is also stored in a separate class. Since the values refer to the corresponding machines, it is also stored in an array list. The first value is the machine ID, the second value is the production time. The values, which are assumed for the production time, are initialized with the help of a manual input of the calculator.

As stated in the expert interviews described above, the production times are estimates based on experience. At this point, the class diagram illustrates which information is used by the calculator to provide an estimation for the times. In addition to the current production order, past production orders are also included in the decision-making process. A list of the work steps and stations and thus the necessary machine assignment can be taken from the work schedule from the production order. Comparable historical orders are then available to the calculator, in which the actual, but also the previously assumed production times can be taken. The associated class ManufacturingOrders_Historical also contains further attributes that can be useful for evaluating the historical data. For example, the employee of the work step, which is stored in the booking system, can be a decision criterion. The attribute machineAllocation provides information about the sequence in which the production took place. Depending on the layout of the production hall, this can provide information about transport routes and associated distortions of the production times in the booking system. The parallelOrders attribute stores the number of orders produced at the same time. A low capacity utilization, for example, is an indication that the production time could have been shorter. In addition to other influencing factors, a deviation between planned and actual values can also be used to improve the expected accuracy of the occupancy times of current production orders. The array list stored as postCalculation as well as the postOfferCalculation is used for this purpose. Intelligent processing of the above-mentioned attributes and influencing factors makes it possible to estimate production costs as precisely as possible.

6. Summary and Outlook

Above all, it becomes clear that the process of quotation costing is currently highly dependent on people and is based in many places on estimations and personal experience. In order to transform quotation costing as a data-supported process, all necessary information and data must be available in used IT system(s). As stated in the very beginning, the interviews also underlined, how many different tools, databases or files are currently used in this process and how time consuming the generation of the prime costs are. To make such an effort for a quotation, not knowing whether the customer will accept it or not, is highly inefficient and risky.

The different criteria used for quotation costing in the case studies are explained and summarized in the data model. Thanks to the UML notation, this data model can be transferred in Business Software or other IT systems to build up a data-bases quotation costing process.

The first finding of the research was, which exact data are needed for the quotation. With the generated results SMEs can check the data model to get an overview, which data they might need to add to their calculation to improve the process and their results. Moreover, the model also shows, which data must be available and can lead companies to the transparency, which process they really need to digitize to get access to data-related information beside the personal experience they mostly rely on so far.

A huge impact for the optimization of the quotation quality is the factor of personal experience and the estimation processes. Especially for important information that are crucial to the quality of the final result the case studies show, that they mainly rely on estimated information. Further research should be conducted here to detail the “Estimator”- Class in the UML diagram, to receive a data-supported estimation. In this area, artificial intelligence applications are particularly conceivable in order to identify similar orders and to draw corresponding conclusions about the current quotation.

References

[1] Abolhassan, F. (Ed.), 2016. Was treibt die Digitalisierung?: Warum an der Cloud kein Weg vorbeiführt. Springer Gabler, Wiesbaden, 138 pp.

[2] Fischer, J., 2019. Datenmanagement: Datenbanken und betriebliche Datenmodellierung. Walter de Gruyter GmbH & Co KG.

[3] Hars, A., 2013. Referenzdatenmodelle: Grundlagen effizienter Datenmodellierung: Schriften zur EDV-orientierten Betriebswirtschaft. Springer-Verlag.

[4] Coenenberg, A.G., Fischer, T.M., Günther, T. (Eds.), 2016. Kostenrechnung und Kostenanalyse. Schäffer Poeschel, Stuttgart.

[5] Albrecht, T., Ulbricht, G., 1931. Die Vorkalkulation Im Kessel- und Apparatebau. Springer, Berlin, Heidelberg.

[6] Kleinaltenkamp, M., Saab, S., 2009. Technischer Vertrieb: Eine praxisorientierte Einführung in das Business-toBusiness-Marketing. Springer, Berlin.

[7] FACTON GmbH, TMG Consultants GmbH, 2015. Herausforderung Angebotskalkulation: Aktuelle Marktumfrage zu Methoden, Organisation, Tools und Handlungsfeldern in Unternehmen der produzierenden Industrie, 56 pp. Accessed 5 January 2020.

[8] Schuh, G., Schmidt, C., 2014. Produktionsmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg, 387 pp.

[9] Obermaier, R., 2017. Industrie 4.0 als unternehmerische Gestaltungsaufgabe. Springer Fachmedien Wiesbaden, Wiesbaden, 320 pp.

[10] Plinke, W., Rese, M., Utzig, B.P., 2015. Industrielle Kostenrechnung. Springer Berlin Heidelberg, Berlin, Heidelberg, 360 pp.

[11] Ehrlenspiel, K., Kiewert, A., Lindemann, U., Mörtl, M., 2014. Kostengünstig Entwickeln und Konstruieren. Springer Berlin Heidelberg, Berlin, Heidelberg, 618 pp.

[12] Knoblach, J., 1999. Beitrag zur rechnerunterstützten verursachungsgerechten Angebotskalkulation von Blechteilen mit Hilfe wissensbasierter Methoden. Dissertation, Bamberg, 178 pp.

[13] Georg Hauschild, 1994. Einwicklung eines Verfahrens zur dynamischen Angebotskalkulation für Unternehmen der Einzel- und Kleinserienfertigung. Dissertation, Aachen.

[14] Denkena, B., Lorenzen, L.-E., Schürmeyer, J., 2009. Rule-based quotation costing of pressure die casting moulds. Prod. Eng. Res. Devel. 3 (1), 87–94.

[15] Liu, W., Yang, C., Zhou, X., 2018. A network quotation framework for customised parts through rough requests. International Journal of Computer Integrated Manufacturing 31 (12), 1220–1234.

[16] Konarsky, M., Leidich, E., Götze, U., 2011. Projektbegleitende Kalkulation komplexer Produkte der Auftragsfertigung. DFX 2011: Proceedings of the 22nd Symposium Design for X, Tutzing nr. Munich, Germany, 11.-12.10.2011 (2), 195–206.

[17] Kadir, A.Z.A., Yusof, Y., Wahab, M.S., 2020. Additive manufacturing cost estimation models—a classification review. Int J Adv Manuf Technol 107 (9-10), 4033–4053.

[18] Ning, F., Shi, Y., Cai, M., Xu, W., Zhang, X., 2020. Manufacturing cost estimation based on a deep-learning method. Journal of Manufacturing Systems 54, 186–195.

[19] Westekemper, M., 2002. Methodik zur Angebotspreisbildung - am Beispiel des Werkzeug- und Formenbaus. Dissertation, 227 pp.

[20] Siepermann, M., 2018. Datenmodell. https://wirtschaftslexikon.gabler.de/definition/datenmodell-28093/version-251730.

[21] Urlaß, F., 2019. Daten- und Prozessmodellierung für Versicherer. Springer Fachmedien Wiesbaden, Wiesbaden, 490 pp.

[22] Shahbaz, Q., 2015. Data Mapping for Data Warehouse Design: Computer Science Reviews and Trends.

[23] Stevens, P., 2006. Using UML. Software engineering with objects and components: The Addison-Wesley object technology series, Harlow.

[24] Gomaa, H. Software modeling and design. UML, use cases, patterns, and software architectures 2011.

Biography

Katharina Berwing M.Sc. (*1991) has been working as a project engineer at the Institute for Industrial Management (FIR) at RWTH Aachen University since 2018. In her current position in the department of Production Management she supports companies in various industries in the fields of selection and implementation of IT systems. She also participates in different research projects.

Prof. Dr.-Ing. Dipl.-Wirt. Ing Günther Schuh (*1958) holds the Chair of Production Systems at the Machine Tool Laboratory (WZL), is a member of the Board of Directors at the Fraunhofer Institute for Production Technology (IPT), Director of the Research Institute for Rationalization e. V. (FIR) at RWTH Aachen University and head of the Production Technology Cluster. He is founder of the Schuh & Co. group of companies based in Aachen, St. Gallen and Atlanta.

Prof. Dr.-Ing. Volker Stich (*1954) has been head of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing. Volker Stich worked for 10 years for the St. Gobain-Automotive Group and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

Smart Products For Smart Production – A Use Case Overview

Günther Schuh[1], Max-Ferdinand Stroh[1], Jan Hicking[1], Sebastian Kremer[1]
[1]FIR – Institute forIndustrial Management at RWTH Aachen University, Aachen, Germany

Abstract

Industry 4.0 is driven by Cyber-Physical Systems and Smart Products. Smart Products provide a value to both its users and its manufacturers in terms of a closer connection to the customer and his data as well as the provided smart services. However, many companies, especially SMEs, struggle with the transformation of their existing product portfolio into smart products. In order to facilitate this process, this paper presents a set of smart product use cases from a manufacturer's perspective. These use cases can guide the definition of a smart product and be used during its architecture development and realization. Initially the paper gives an introduction in the field of smart products. After that the research results, based on case-study research, are presented. This includes the methodological approach, the case-study data collection and analysis. Finally, a set of use cases, their definitions and components are presented and highlighted from the perspective of a smart product manufacturer.

Keywords

Industry 4.0; Smart Products; Digitalization; Smart Machines; Product-Service-Systems

1. Introduction

The digitalization found its way into the manufacturing environment. Even traditional industries such as mechanical engineering are affected by this trend [1]. The products of mechanical engineering, tools and machines, which previously contained only mechanical parts are becoming smart and connected. In other words, they are enhanced with electronic and digital components such as sensors and microprocessors [2]. The result are smart products. They are based on cyber-physical systems and consist of both physical and digital components [1]. They form the foundation of industry 4.0 and by that, the establishment of collaborative networks [3,4].

Smart products can be used to generate a competitive advantage by improving the services provided to their users [5]. Smart products are an exceptional opportunity of getting a data based, detailed understanding of a company's customer and how the product is actually used. This allows great insight into what future products might look like and which features need to be updated. Smart services can be tailored and continuously improved to serve a customer's need. [6,2,5]

However, to gain those valuable insights, a well architected digital environment needs to be built around a product, finally making it and its insights smart [2,5]. Many companies, especially SMEs struggle with this challenge, as they are lacking the internal knowledge for developing such complex digital products and architectures and even expect a business loss if they fail to place a smart product on the market [6,7].

In order to facilitate the process of developing a smart product as well as helping the business decisions for smart products, an overview of its possible applications in terms of use cases is needed. In the long run those

DOI: https://doi.org/10.15488/12146

use case can serve as a basis for the development of the digital architecture of smart products. This is the goal of the underlying research project BlueSAM, which provides SMEs with blueprints for the development of smart products architectures.

For this reason, this paper presents a set of use cases for smart products from a smart product manufacturer's perspective. First of all, the term smart product is defined in the context of this paper. Subsequently the overarching research project and the envisioned smart product development process for SMEs is described. After that the process of case study research is explained and applied. This leads to an empirically derived set of use cases for smart products, that are validated by industry experts and literature.

2. State of the art

Smart products are in the limelight of Industry 4.0 and one of the main topics among the smart factory, smart logistics and smart development [6]. Smart Products, their definitions and differentiations have been widely covered in scientific literature, such as [8,2] and the authors' previous works [9,1,10]. Based on HICKING'S definition in [1,10] the authors define a smart product as "a product, which consists of both a physical and a digital component. They create value for both, its user (mostly the customer) as well its manufacturer. For users a smart product's main added value is to provide smart services. For the manufacturer it is the opportunity to learn from its newly generated usage data". Smart services are a data based combination of both digital and physical services provided by smart products [11].

In mechanical engineering there are two main fields of application for smart products: the use of smart products in the own production and the sale of smart products as well as accompanying services [12]. This paper will focus on the perspective of a smart product manufacturer. Examples for smart products in manufacturing are smart industrial air compressors, software applications for a mobile steering of machines or connected machines which are capable of gathering and analyzing production data in real-time [13,14].

The applications of smart products are only limited by imagination. However, several publications have offered overviews of existing smart product use cases or described specific use cases. Examples are PORTER & HEPPELMANN [2,5], HERTERICH ET AL. [15,16], ABRAMOVICI [17] or MACHCHHAR ET AL. [8]. As this work employs smart product use cases derived from practical applications by Case Study Research [18], the aforementioned works – among others - are applied in their literature based verification process (see section 4.3). This results in the use case overview presented in this paper.

3. Methodology and focus of this paper

As mentioned before the results of this paper are part of an overall research project assisting SMEs in the development of their smart products' digital architecture and infrastructure. This will be with a use case based, smart product reference architecture with focus on SMEs. From this reference architecture, for every use case, blueprints for a smart product architecture will be derived. The development of the reference architecture is based on the process by KRCMAR ET AL. [19]. This will lead to an overall process for the facilitated development of smart products based on blueprints. The process is drafted in Figure 1. Step one is the selection of smart product use cases in the beginning of the development process. It will give SMEs an inspiration into which applications can be realized with their future smart products. Secondly, architecture blueprints will be selected based on the selected use cases. In the third step, the blueprints will be specified and customized into a company-specific architecture. After that the, the process allows the delegation of development tasks among partner companies and external service providers marking the Co-Development phase. This fourth step considers, that most SMEs do not have all capabilities for smart product development in house and therefore need help in form of a Co-Development guideline for partnering up with and managing external partners.

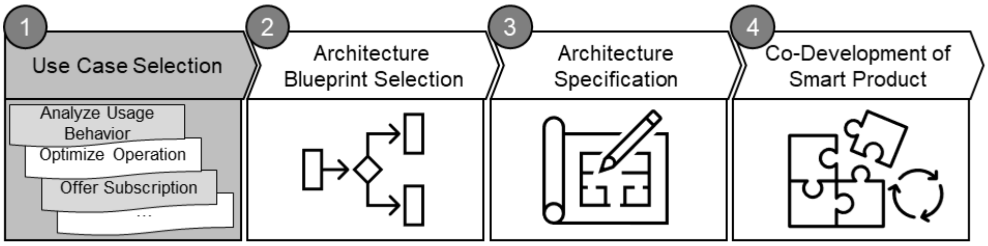


Figure 1: Overall project approach (grey: focus of this paper)

This paper focuses on step one of the envisioned development process, providing a list of smart product use cases based on a smart product manufacturer's perspective. To be up to date, the approach of case study research by EISENHARDT [18,20] was chosen to derive the use cases. It provides use cases derived from empirical evidence, validated with existing literature [18,20,21]. Furthermore, the selected research approach provides valuable insights into the realization of the use cases at the case study partners that could not be achieved by solely focusing on theoretical evidence and existing use case models.

In case study research, theories or hypotheses are evaluated with the help of so called "case studies", that represent empirical data collections from multiple data sources. During the research case studies are collected and then analyzed individually as well as between each other. The identified hypotheses are then developed and enhanced from case to case. After that the derived theories are then verified based on a literature analysis. After 4 to 10 case studies a saturation of the built theory should be reached, meaning, there are no more additions to it [18,20,22,21]. Case study research has become a popular research methodology beyond its origins in social sciences as it delivers an opportunity to analyze theoretical situations where a statistical approach is not feasible or less effective [20,21].

4. Smart Product Use Cases derived from Case Studies

4.1 Selected Case Studies and saturation

In compliance with case study research the set of case studies was very carefully selected in order to lead to valid results. Therefore the selection of use cases was based on theory not on sampling [18,21]. Case study research aims at a replication of results from case to case not on a statistical sampling [18]. Table 1 gives an overview of the case studies selected for this research. The case study companies were selected with the goal to deep dive on use cases analyzing usage data as well as providing smart services to customers. The case studies were selected by company size, measured in number of employees and revenue, the type of their business relationship (business to customer (B2C) or business to business (B2B)), the complexity of their smart product (machine or assembly group), the current lifecycle phase it is in (in use or in development) and their type of company (smart product manufacturer or smart product service provider). The selected set of case studies is both not theoretically identical but also not a random sample of companies. For the company type, next to manufacturers, service providers were chosen on purpose as they offer several smart products or cater to many different customers and thus have a wide knowledge of different applications and use cases for smart products.

Table 1 : Overview of selected case studies

No.	Employees	Revenue	Business type	Product complexity	Lifecycle phase	Company Type
1	> 500	< 100 M. €	B2C, B2B	Machine	Usage	Manufacturer
2	> 1000	< 250 M. €	B2B	Machine	Usage	Manufacturer
3	> 6500	< 1,800 M. €	B2B	Machine	Usage	Manufacturer
4	> 2000	< 450 M. €	B2B	Machine	Usage	Manufacturer
5	> 1500	< 3,000 M. €	B2C, B2B	Machine	Development	Manufacturer
6	> 150	< 50 M. €	B2C, B2B	Machine	Development	Manufacturer
7	> 15	< 0,5 M. €	B2C, B2B	Assembly group	Usage	Service Provider
8	> 6000	< 1,500 M.€	B2B	Machine	Usage	Service Provider
9	> 10	< 3 M. €	B2C, B2B	Assembly group	Development	Service Provider
10	> 10000	< 2,000 M. €	B2B	Assembly group	Usage	Manufacturer

For every case study an individual company was selected. Different data sources such as expert interviews performed by the authors of this paper, publications, presentations or product tests were used in the case studies' analysis. During the data analysis, smart product use cases were derived from the collected case studies and compared to the a-priori use case model. Thusly the smart product use case model evolved during the research. As Figure 2 is showing, after seven case studies, no more new use cases were identified and added to the model. This means saturation is reached.

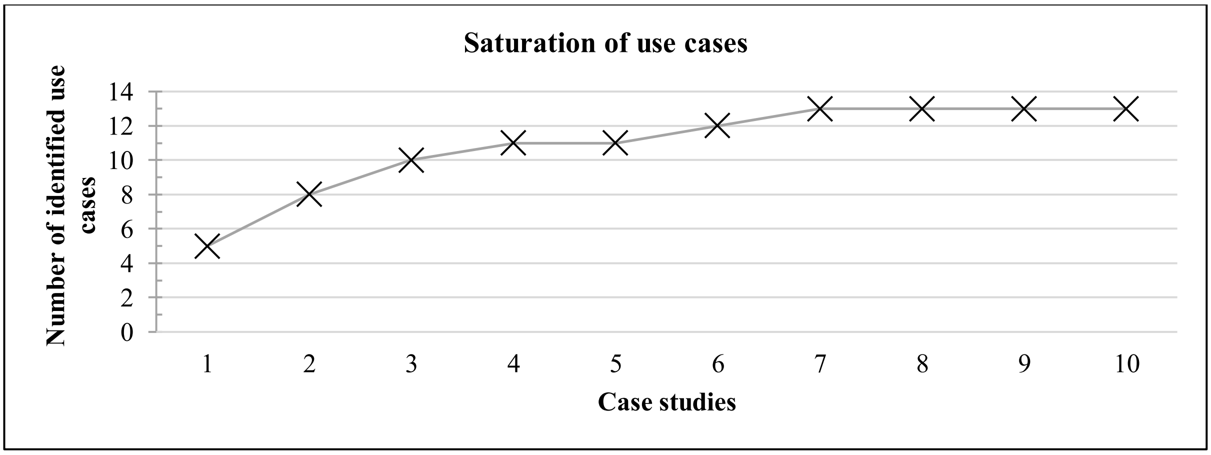


Figure 2: Saturation curve of case study analysis

4.2 Identified Smart Product use cases

In Total 13 use cases were derived which can be distinguished into two main categories. They are shown in Table 2. The first category is focused on use cases focusing on usage and field data. The second category is focused on providing smart services for the smart product's users. The use cases are explained in detail in the following subsections. The descriptions of the use cases represent the aggregated and anonymized findings from the case study analysis. All of the use cases were verified during expert interviews with the user committee of the underlying research project and literature about smart product use cases.

Table 2: Overview of derived smart product use cases

Use Cases with focus on usage data	Use Cases with focus on smart services
Analyze usage behavior	Offer Condition Monitoring
Derive new products or services	Provide Predictive Maintenance

Improve existing products or services	Offer data analytics
Update products or services	Assist operation
Upgrade products or services	Optimize operation
Offer Subscription	Deliver consumables or supplies
	Create digital product image

4.3 Description of the use cases derived from the analyzed case studies

4.3.1 Derive new products or services

The data collected from smart products can be used to identify opportunities for new products and services. This means identifying completely new product ideas, missing product functions in existing smart products or changing the configuration of features in a smart product [23,15,24]. This means omitting unused functions or combining them into new functions. To do so, the existing product configuration is measured against the identified customer needs from the usage behavior analysis as well as the product usage and linked to product functions [25]. From here missing functions can be identified and evaluated for new products [25].

4.3.2 Improve existing products or services

Next to deriving new products and services, existing products or services can be improved with usage data from smart products. Though this use case is very similar to *derive new products or services* it is mentioned separately. The matter of differentiation for the authors is the improvement of existing functions versus the identification of missing ones, which leads to new products or services. In this case the focus of the analysis is set on the existing functions and features and if they can be improved, e.g. for better user experience or performance [26,15,24,2].

4.3.3 Update products or services

The *update products or services* use case means improving or enhancing the existing functions smart product by exchanging physical or digital components [27]. To do so, an update needs to be ready for deployment from the manufacturer. This may be an improvement of software such as a change in the user interface but may also include the exchange of a physical component [24,2]. Before that the update needs to be authorized by the user. If necessary, a downtime is scheduled for the smart product. Depending on service level agreements, an exchange product may be delivered on site during the smart products update time [2,5].

4.3.4 Upgrade products or services

Upgrading a product or a service, in contrast to an update, refers to deploying new digital or physical functions to a smart product, including an entire exchange of the product [28,14,27]. The process of upgrading a smart product is analogous to the update process. Meaning an upgrade is ready for deployment and will be executed similarly to the update (see section 4.3.3). To identify the need for an upgrade the usage of the smart product is monitored against predefined performance KPIs. Once a threshold is exceeded, the current product use is evaluated. This could mean using a smart product at its upper or lower performance limit triggers an exchange against a smaller or bigger product to better fit the customer's usage behavior [14]. The use of upgrades extends the lifecycle of products and the level of customer satisfaction by continuously satisfying the customer needs [28]. Furthermore it also allows a continuous evaluation and change of the smart product based on its usage [28].

4.3.5 Analyze usage behavior

As the use cases are used in the conceptualization phase during the development of a smart product, a detailed yet easy to understand visualization is needed. Therefore, in addition to the use case description every use case is visualized in a UML Use case diagram [29]. It is a type of visualization that is easy to understand, solution agnostic, and yet leaves enough space for interpretation as the design of a smart product is evolving. Nevertheless, it contains the main components that are needed to realize a smart product. The use case diagram for *analyze usage behavior* is shown in Figure 3.

This use case focuses on understanding the user's behavior by analyzing usage data collected by the smart products [30,15,2]. It builds the foundation for better understanding how and why customers are using the smart product and their specific needs [30,15,31,8,2]. It means collecting the usage data such as frequency, duration, and location of usage as well as the used features from the smart product and storing it within the product cloud [30,8,2]. In addition to that, user feedback on the product itself or certain features is added to the analysis [24]. By analyzing the usage behavior over a certain period of time a behavioral profile of the user can be derived [24,2]. The collected data is then matched with customer specific data as well as data about the customer' s segment [30,2]. In addition to that the customer's usage profiles can be compared to one another determining certain usage patterns and customer needs [2,25]. In order to collect usage data, this use case will collect the usage data based on the use case *offer condition monitoring* (section 4.3.7) as well as customer feedback based on *assist operation* (section 4.3.10).

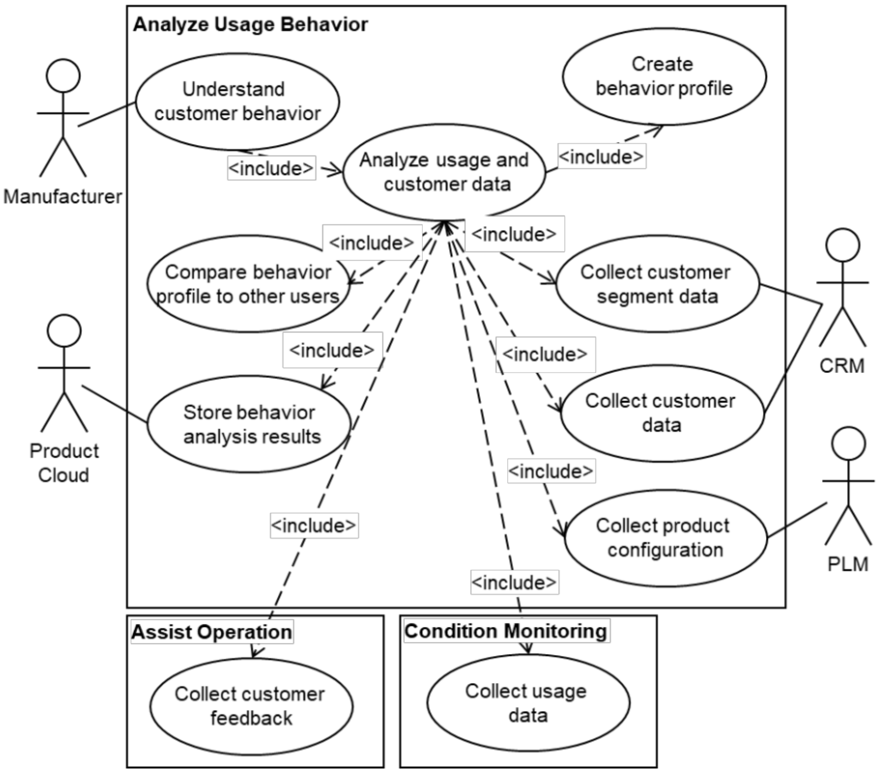


Figure 3: Use Case diagram for *analyze usage behavior*

4.3.6 Offer Subscription

Offering a subscription business model means delivering continuous value improvement for a fixed fee [32,33]. The online streaming service Netflix – with a fixed price rate and ever growing offering – is an example for a subscription based business model from a B2C context [34]. Moving from a transaction based customer interaction to an outcome based customer interaction is becoming increasingly popular in the B2B context and is enabled by smart products [32]. The delivery of value is often measured in product performance [35,15]. The improvement in performance is achieved by analyzing the usage behavior to identify opportunities for process or product improvements for the specific customer use case [35,2].

4.3.7 Offer Condition Monitoring

The use case *offer condition monitoring* consists of visualizing a products condition as well as providing the user with alerts in case predefined thresholds are exceeded or a set of rules triggers it. This means that the user is allowed to define alerts and rules [26,2,5]. To execute the rules the smart product collects data from different sources and aggregates them [15]. The data may be from the product's environment such as brightness, humidity (environmental data), from within the product such as internal sensor values or errors (product data) or regarding the process the product is in (process data) [8]. The later visualized or monitored values may be specific values such as the vibration within a motor or KPIs aggregating different data streams. This could be the asset health or for example the smart product's OEE [26]. The condition monitoring use case serves as a base for many of the identified use cases. This use case can also be applied to the smart product's manufacturer monitoring certain KPIs for the smart product during its usage [26,35].

4.3.8 Provide Predictive Maintenance

Predictive Maintenance aims at providing the user of the smart product with an interruption free usage period. It means identifying the need for maintenance and scheduling it before the smart product breaks down unexpectedly, avoiding unforeseen downtime [36]. In the case of an identified maintenance need a planned maintenance will be scheduled automatically to ensure maximum performance of the smart product [2,14]. To do so the remaining usable life of components or the whole smart product is calculated and monitored. This is done by connecting the as is data from the smart product with historical data and historical maintenance cases [8,5]. Based on the identified issues, maintenance measures are selected and a maintenance is scheduled. Depending on the situation, the maintenance can be executed remotely [26]. As the component to be exchanged is known in advance the maintenance personnel can bring the right equipment and plan the maintenance process accordingly [2,14]. If needed an exchange product can be provided during the maintenance period [37].

4.3.9 Offer data analytics

This use case allows the smart product's user to individually deploy data analytics models on the smart product using its field and product data. This means selecting from a set of predefined analytics models that can be applied to the smart product and its environment [23]. These can be anomaly detection, assistance for teach in of sensor values, vibration and temperature analytics, etc. [5]. It includes external data sources such as business systems as well as storing and exporting the analytics results [2]. Offering such functions may have a positive effect on the perceived value of the smart product.

4.3.10 Assist operation

Assisting the operation refers to helping the smart product's user to operate it. This means the smart product automatically guides its user. It monitors the way it is used and provides helpful information such as manuals, warning about dangerous or unintended maluses of the product and prevents the user from making mistakes [15,16,2]. It will allow remote assistance and remote control as well as collect feedback from the users via its digital interface [35,2].

4.3.11 Optimize operation

Optimize operation refers to helping the smart product's user to operate it at an optimal state. This is done via recommendations for the user to improve the product's performance or lifetime [2,14]. The recommendations can be drawn from a predefined set of recommendations. The recommendations are selected on the base of an optimization model constantly comparing the current operation parameters to an optimal state [26,15,2]. The smart product's data can be enriched with external data sources like the production schedule from an MES or ERP software [2].

4.3.12 Deliver consumables or supplies

In this use case the smart product watches the usage of supplies or consumables during its use. Based on the usage it can estimate the remaining time until new supplies will be needed. Being connected to the stock management system at the company or the supplier it can order material just in time to not run out of stock [38,15,14]. This may also lead to a more detailed usage understanding for the product's manufacturer also offering an additional value stream in case the supplies are directly sold by him.

4.3.13 Create a digital product image

A digital product image gives an overview of historic data from the smart product based on a continuous data flow [39,40]. In the presented set of use cases, the term *digital product image* is used synonymously for *digital shadow* as well as *digital product passport* [39]. It stores all of the historical product data and allows the user to explore it from different views and contexts. Such views may be the manufacturing history or the maintenance history of the smart product. Additionally the data can be exported to other systems or be analyzed in a different context [39,41,40].

5. Summary and outlook

In this paper a set of smart product use cases for developing a smart product are presented. They are described from the perspective of a smart product manufacturer with industrial application. The overarching smart product development process is explained. The use cases are derived with case study research. The methodology as well as the selection and saturation of case studies and use cases is explained. After that, the validated use cases are described in detail and an exemplary use case diagram is shown.

Future research will focus on building blueprints for easily applying the use cases during the smart product development process as well as creating a methodology for a focused co-development of the smart products.

Acknowledgements

The Cornet project 303 EN of the Research Association FIR e. V. at the RWTH Aachen University is funded via the AiF within the framework of the programme for the funding of cooperative industrial research (IGF) by the Federal Ministry for Economic Affairs and Climate Action (BMWK) on the basis of a decision by the German Bundestag.

References

[1] Hicking, J., Zeller, V., Schuh, G., 2018. Goal-Oriented Approach to Enable New Business Models for SME Using Smart Products, in: Chiabert, P., Bouras, A., Noël, F., Ríos, J. (Eds.), Product Lifecycle Management to Support Industry 4.0. 15th IFIP WG 5.1 International Conference, PLM 2018, Turin, Italy, July 2-4, 2018, Proceedings. Springer International Publishing, Cham, pp. 147–158.

[2] Porter, M.E., Heppelmann, J.E., 2014. How Smart, Connected Products Are Transforming Competition. Harvard Business Review (9), 1–23.

[3] Abramovici, M., Göbel, J.C., Savarino, P., 2016. Virtual Twins as Integrative Components of Smart Products, in: Harik, R., Rivest, L., Bernard, A., Eynard, B., Bouras, A. (Eds.), Product Lifecycle Management for Digital Transformation of Industries, vol. 492. Springer International Publishing, Cham, pp. 217–226.

[4] Camarinha-Matos, L.M., Fornasiero, R., Afsarmanesh, H., 2017. Collaborative Networks as a Core Enabler of Industry 4.0, in: Camarinha-Matos, L.M., Afsarmanesh, H., Fornasiero, R. (Eds.), Collaboration in a Data-Rich World, vol. 506. Springer International Publishing, Cham, pp. 3–17.

[5] Porter, M.E., Heppelmann, J.E., 2015. How Smart, Connected Products Are Transforming Companies. Harvard Business Review (10), 96–112.

[6] Abramovici, M., Gebus, P., Savarino, P., 2018. Engineering smarter Produkte und Services: Plattform Industrie 4.0 STUDIE, München, 46 pp.

[7] Lünnemann, P., Wang, W.M., Lindow, K., 2019. Smart Industrial Products: Smarte Produkte und ihr Einfluss auf Geschäftsmodelle, Zusammenarbeit, Portfolios und Infrastrukturen, 66 pp.

[8] Machchhar, R.J., Toller, C.N.K., Bertoni, A., Bertoni, M., 2022. Data-driven value creation in Smart Product-Service System design: State-of-the-art and research directions. Computers in Industry 137, 103606.

[9] Hicking, J., 2020. Spezifikation von intelligenten Produkten im Maschinenbau. Apprimus, Aachen, 384 pp.

[10] Stroh, M.-F., Hicking, J., Stich, V., 2021. Smartifizierung von Maschinenbauprodukten mittels einer zielorientierten Methode, in: Meinhardt, S., Wortmann, F. (Eds.), IoT – Best Practices. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 259–275.

[11] Kampker, A., Frank, J., Jussen, P., 2017. Digitale Vernetzung im Service. WIST 46 (5), 4–11.

[12] Bauer, W., Schlund, S., Marrenbach, D., Ganschar, O., 2014. Industrie 4.0: Volkswirtschaftliches Potenzial für Deutschland, 46 pp.

[13] Kinkel, S., Rahn, J., Rieder, B., Lerch, C., Jäger, A., 2016. Digital-vernetztes Denken in der Produktion, Frankfurt am Main, 88 pp.

[14] Urbach, N., Röglinger, M., 2019. Digitalization Cases. Springer International Publishing, Cham, 429 pp.

[15] Herterich, M.M., Uebernickel, F., Brenner, W., 2015. The Impact of Cyber-physical Systems on Industrial Services in Manufacturing. Procedia CIRP 30, 323–328.

[16] Herterich, M.M., Uebernickel, F., Brenner, W., 2016. Industrielle Dienstleistungen 4.0. Springer Fachmedien Wiesbaden, Wiesbaden.

[17] Abramovici, M., 2014. Smart Products, in: Produ, T.I.A.f., Laperrière, L., Reinhart, G. (Eds.), CIRP Encyclopedia of Production Engineering, vol. 59. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–5.

[18] Eisenhardt, K., 1989. Building Theories from Case Study Research. The Academy of Management Review 14 (4), 532–550.

[19] Krcmar, H., Reidt, A., Duchon, M., 2017. Erstellung einer Referenzarchitektur anhand von individuellen Unternehmensanforderungen, in: Bullinger-Hoffmann, A.C. (Ed.), S-CPS: Ressourcen-Cockpit für Sozio-Cyber-Physische Systeme. Abschlussveröffentlichung. aw&I - Wissenschaft und Praxis, Chemnitz, pp. 23–42.

[20] Eisenhardt, K., Graebner, M., 2007. Theory Building From Cases: Opportunities and Challenges. Academy of Management Journal (50), 25–32.

[21] Yin, R.K., 2018. Case study research and applications: Design and methods, Sixth edition ed. Sage, Los Angeles, 319 pages ;

[22] Gassmann, O., 1999. Praxisnähe mit Fallstudienforschung: Nutzen für das Technologiemanagement ist gegeben. Wissenschaftsmanagement (3), 11-16.

[23] Alcayaga, A., Wiener, M., Hansen, E.G., 2019. Towards a framework of smart-circular systems: An integrative literature review. Journal of Cleaner Production 221, 622–634.

[24] Pfeifer, T., Schmitt, R. (Eds.), 2021. Masing Handbuch Qualitätsmanagement, 7., überarbeitete Auflage ed. Hanser, München, 1091 pp.

[25] Riesener, M., Dölle, C., Schuh, G., Becker, A. Framework for the Continuous Increase of Product Performance by Analyzing Product Usage Data: IEEM2019 : 15-18 Dec, Macau, in: .

[26] Basselot, V., Berger, T., Sallez, Y., 2017. Active Monitoring of a Product: A Way to Solve the "Lack of Information" Issue in the Use Phase, in: Borangiu, T., Trentesaux, D., Thomas, A., Leitão, P., Oliveira, J.B. (Eds.), Service Orientation in Holonic and Multi-Agent Manufacturing, vol. 694. Springer International Publishing, Cham, pp. 337–346.

[27] Zeller, V., 2018. Releasemanagement für digitale Produkte: Vortrag im Rahmen der mündlichen Doktorprüfung von Dipl.-Inform. Violett Zeller.

[28] Khan, M.A., Wuest, T., 2018. Towards a framework to design upgradable product service systems. Procedia CIRP 78, 400–405.

[29] Object Management Group, 2017. OMG Unified Modeling Language (OMG UML): Version 2.5.1, 796 pp.

[30] Exner, K., Smolka, E., Blüher, T., Stark, R., 2019. A method to design Smart Services based on information categorization of industrial use cases. Procedia CIRP 83, 77–82.

[31] Hollauer, C., Shalumov, B., Wilberg, J., Omer, M., 2018. Graph Databases for Exploiting Use Phase Data in Product-Service-System Development: A Methodology to Support Implementation, in: Proceedings of the DESIGN 2018 15th International Design Conference. 15th International Design Conference. May, 21-24, 2018. Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia; The Design Society, Glasgow, UK, pp. 1571–1582.

[32] Günther Schuh, Lucas Wenger, Volker Stich, Jan Hicking, Jonas Gailus. Outcome Economy: Subscription Business Models in Machinery and Plant Engineering.

[33] Tzuo, T., Weisert, G., 2018. Subscribed: Why the subscription model will be your company's future - and what to do about it. Portfolio/Penguin, 256 pp.

[34] Costa, C.D., 2020. How Data Science is Boosting Netflix - Towards Data Science. Towards Data Science, April 19.

[35] Herterich, M.M., Eck, A., Uebernickel, F., 2016. Exploring How Digital Products Enable Service Innovation: An Affordance Perspective. Research Papers 156. 24. European Conference on Information Systems (ECIS), İstanbul,Turkey, 2016 4, Istanbul (Türkei), 18 pp. http://aisel.aisnet.org/ecis2016_rp/156. Accessed 17 December 2017.

[36] Werner, A., Zimmermann, N., Lentes, J., 2019. Approach for a Holistic Predictive Maintenance Strategy by Incorporating a Digital Twin. Procedia Manufacturing 39, 1743–1751.

[37] Gassmann, O., Frankenberger, K., Csik, M., 2017. Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler Business Model Navigator, 2., überarbeitete und erweiterte Auflage ed. Hanser, München, 1412 pp.

[38] Anke, J., Wellsandt, S., Thoben, K.-D., 2018. Modelling of a Smart Service for Consumables Replenishment. 17:1–21 Pages / Enterprise Modelling and Information Systems Architectures (EMISAJ), Vol 13 (2018).

[39] Becker, F., Bibow, P., Dalibor, M., Gannouni, A., Hahn, V., Hopmann, C., Jarke, M., Koren, I., Kröger, M., Lipp, J., Maibaum, J., Michael, J., Rumpe, B., Sapel, P., Schäfer, N., Schmitz, G.J., Schuh, G., Wortmann, A., 2021. A Conceptual Model for Digital Shadows in Industry and Its Application, in: Ghose, A., Horkoff, J., Silva Souza, V.E., Parsons, J., Evermann, J. (Eds.), Conceptual Modeling. 40th International Conference, ER 2021, Virtual Event, October 18–21, 2021, Proceedings, vol. 13011, 1st ed. 2021 ed. Springer International Publishing; Imprint Springer, Cham, pp. 271–281.

[40] Schuh, G., Gützlaff, A., Sauermann, F., Maibaum, J., 2020. Digital Shadows as an Enabler for the Internet of Production, in: Lalic, B., Majstorovic, V., Marjanovic, U., Cieminski, G. von, Romero, D. (Eds.), Advances in Production Management Systems. The Path to Digital Transformation and Innovation of Production Management Systems, vol. 591. Springer International Publishing, Cham, pp. 179–186.

[41] Schacht, M., Schuh, G., Frank, J., 2021. The digital shadow of customers in the manufacturing industry, in: 2021 8th Swiss Conference on Data Science (SDS). 2021 8th Swiss Conference on Data Science (SDS), Lucerne, Switzerland. 09.06.2021 - 09.06.2021. IEEE, pp. 52–53.

Biography

Günther Schuh (*1958) is head of the chair of Production Systems (WZL-PS) at RWTH Aachen University and member of the directorates of the Machine Tool Laboratory (WZL) at the RWTH, Fraunhofer Institute for Production Technology (IPT) and Director of the FIR at RWTH Aachen University.

Max-Ferdinand Stroh (*1991) is a researcher at FIR at RWTH Aachen University since 2017 in the department Information Management. He is head of the department Information Management. His scientific work is focused on the practical application of AI, smart products, and IT-OT-Integration.

Dr.-Ing. Jan Hicking (*1991) has been head of the division Information Management at FIR at RWTH Aachen University since 2020. Starting in 2016, he received his Ph.D. in 2020 in the field of smart products. As deputy head of the division, he is responsible for multifaceted consulting and research projects.

Sebastian Kremer (*1993) Is leader of the group Information Technology Management at FIR at RWTH Aachen University. His work is focused on digital technologies and their applications such as smart products and event-driven architectures.

3rd Conference on Production Systems and Logistics

Process Chain of Injection Moulding And Additive Manufacturing For Hybrid Parts

Johann Schorzmann[1], Daniel Pezold[1] , Jan Kemnitzer[2],
Alexander Kalusche[3], Frank Döpper[1,2]
[1]*University of Bayreuth – Chair of Remanufacturing and Manufacturing Technology, Bayreuth, Germany*
[2]*Fraunhofer Institute for Manufacturing Technology and Automation IPA, Bayreuth, Germany*
[3]*Acad Group GmbH, Heilsbronn, Germany*

Abstract

To achieve lightweight properties, a process for the local reinforcement of injection-moulded parts using additive manufactured continuous fibre reinforced inserts was developed. The process is based on the additive manufacturing (material extrusion) of semi-finished products (inserts) made of pre-impregnated continuous fibres, which are inserted in a defined position of the injection mould and then over-moulded with a compatible polymer matrix. Due to the manufacturing method of the inserts, a material, form and frictional bonding between the fibres and the polymer matrix can be realised. By choosing a specific positioning of the inserts resp. the continuous fibres, one can achieve a significant lightweight construction potential and the targeted elimination of process-related weak points such as weld lines of the injection-moulded parts.

A virtual development process using digital product development tools were applied for the construction of a concrete application example. A combination of topology optimisation and finite element analysis (FEA) was used to determine the load- and material-optimised design. For the simulation and optimisation of hybrid parts a new method of virtual development and definition of material models for additively manufactured components was developed. The validation of the process chain for hybrid parts was carried out using a test setup that represents real load situations of an automotive part. The technical analysis of the hybrid part showed a weight saving of 19.5% compared with the reference part. Regarding the critical load case (load from above), a 38.1% lower deformation was achieved. The specified maximum load and deformation limits were maintained in the use case. In addition, in the weld line area malfunction was avoided by the continuous fibre-reinforced insert.

Keywords

Injection moulding; additive manufacturing; hybrid part; continuous fibre reinforced inserts; topology optimisation

1. Introduction

Over decades, numerous manufacturing processes have been developed that enable the processing of polymers. Injection moulding has established itself as a standard process in industrial applications for the production of polymer parts. Due to the high level of existing know-how and the optimal control of the processes in the field of injection moulding, this manufacturing process enables the high-quality and cost-

DOI: https://doi.org/10.15488/12129

effective production of polymer parts in large quantities (large-scale production). The wide range of available materials also makes injection moulding interesting for the industry. However, the large number of advantages is offset by a number of disadvantages. The strengths and stiffness's when using thermoplastics are low in relation to metal materials and are also stress-dependent [1]. Due to these circumstances, these injection-moulded parts are not suitable for high stresses, which means that the weight advantage of metal materials is no longer decisive.

To leverage more of the lightweight potential of injection-moulded parts, it is necessary to improve their mechanical properties. Technically feasible approaches already exist, in which unfilled plastics are substituted in the injection moulding process by short (0.1 – 1.0 mm) or long (1.0 – 50.0 mm) glass fibre reinforced polymers. However, the use of significantly better reinforcing continuous fibres (endless) is not feasible here, as these fillers cannot be processed in injection moulding. Therefore, there is still a need for research in the case of local stress and simultaneously required minimum weight. [1] For these requirements, however, one alternative is to integrate continuous fibre into an injection-moulded part using an additional manufacturing process. In this case, additive manufacturing can be used to significantly improve the mechanical properties of polymer parts and utilizing the geometrical freedom to locally reinforce the polymer part [2-8].

Accordingly, the aim of this scientific work is to develop a concept to improve the mechanical capabilities and at the same time save material and reduce weight of an injection moulded part by inserting an additive manufactured continuous fibre reinforced part. For this purpose, a virtual product development process is applied.

2. State of the Art

Insert technology is one way to open up further lightweight potential and at the same time improve the mechanical properties of an injection-moulded part. Here, the designing engineer can react to the local stresses and reinforce the stress hotspots in the injection-moulded part by locally positioned inserts. In polymer processing techniques usually, the inserts are made of metallic materials. Using metal inserts can create a friction or form bond and for material bonding a post processing step has to be applied. However, when the part volume is small, it is difficult to transfer the load to the metal. Following the example of long glass fibre reinforced injection moulding, inserts made of continuous fibre reinforced polymer composites offer new manufacturing possibilities. Thus, mechanical properties can be significantly increased at the areas subjected to the highest local stresses. For example, local reinforcement offers the possibility of reducing ribbing, wall thicknesses, and the long glass fibre filler content, which leads to further material and weight savings [1]. To realize the material and weight savings, continuous fibres can be integrated into an injection-moulded part via an additional upstream manufacturing process. The additive manufacturing technologies offer new approaches for this purpose.

The additive manufacturing process **of** material extrusion (MEX), which uses filaments as semi-finished products for producing parts, is increasingly developing from a manufacturing process for prototypes to a manufacturing process for end use parts. Similar to other polymer processing methods, mechanical properties of the manufactured parts can be further enhanced by MEX through the addition of fillers such as nanomaterials, particles or short fibres to the used filament. These composite filaments are characterised by high mechanical properties at low weight and very good functionality. However, the mechanical properties of short fibre reinforced composites produced by additive manufacturing are still more anisotropic with composites produced by conventional methods, depending on the orientation of the additive manufactured part and the processing parameters of the additive machine. For the production of continuous fibre reinforced thermoplastic parts, MEX promises to be an alternative to conventional process chains, since no cost-intensive equipment, such as tooling or autoclave, is required [9, 10]. In general, the additive manufacturing process MEX for producing continuous fibre reinforced polymer parts uses a multi-material approach. Two filaments (continuous fibre reinforced filament with polymer matrix and an unfilled polymer filament) are heated up and extruded layer by layer in order to manufacture a continuous fibre reinforced polymer part (see Figure 1).

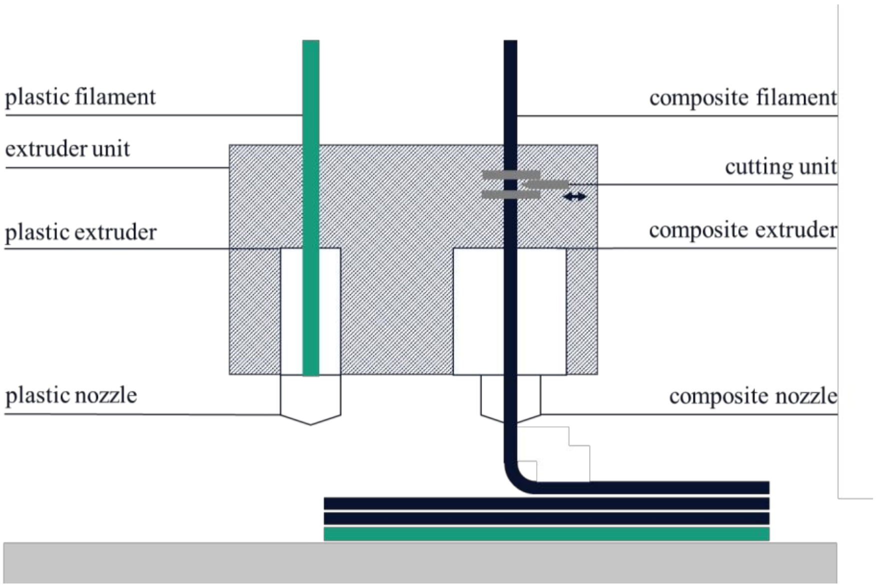


Figure 1: Schematic structure of the Additive Fusion Technology [11]

Therefore, the geometry of the additive manufactured continuous fibre reinforced polymer inserts can be individually and digitally adapted to the exposure path. In general, when processing the continuous fibre filament, their dimensions must be taken into account, since their stiffness means that they cannot be processed in the MEX in the same way as the unfilled polymer filaments because of the lower bending radius in comparison to the unfilled polymer filaments.[11]

9T Labs' Additive Fusion Technology (AFT) uses the strategy of upstream co-extrusion process for composite filament production and dual extrusion method in the manufacturing process (Figure 1). The composite filament's fibre volume content of 60% is very high compared to others available on the market. [11, 12]

3. Methodology

The objective was to improve a reference part from the automotive industry: armrest. To increase the strength combined with a simultaneous reduction in weight, further development in design, simulation and production technology were necessary. The standard reference part was made of a long glass fibre reinforced thermoplastic using an injection moulding process. The approach pursued to increase strength was the use of continuous fibres, which could not previously be introduced or processed in the injection moulding process. The continuous fibres were integrated in an upstream production stage. Therefore, in order to achieve an improvement in strength, locally additively manufactured continuous fibre reinforced polymer inserts were implemented in the reference part. Additive manufacturing could be used to individually manufacture filigree structures from pre-stretched continuous fibre reinforced polymer filament with a rough surface without the need for tools, which was not possible using previous manufacturing technologies. The rough surface in case of the high fibre content is used for better bonding in injection mould and the continuous fibre reinforced insert. The simultaneous weight savings were achieved by using virtual product development tools to design the hybrid part. By identifying the load flow in the reference part, the part could be strengthened at highly loaded locations and material saved at less loaded locations (see Figure 2)

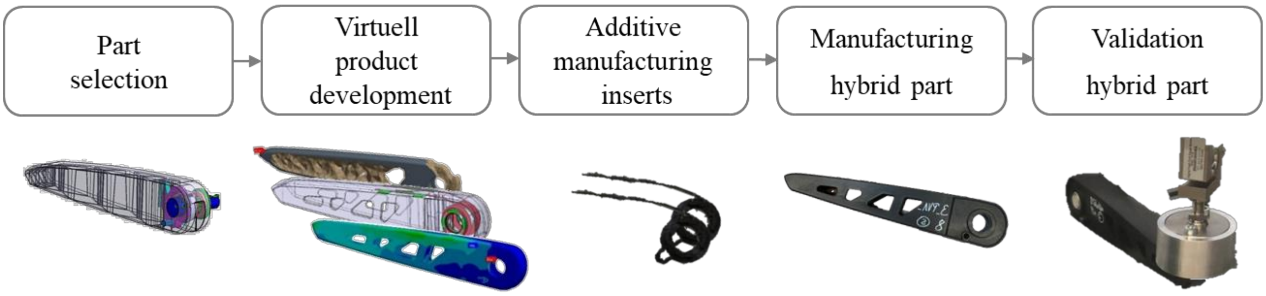


Figure 2: Procedure for the developed hybrid part

In the first step of the process chain for hybrid parts, the boundary conditions of the reference part were defined, which reflected the conditions of use in reality. In particular, the load situation and the mounting conditions of the reference part were to be mentioned as boundary conditions. Based on these boundary conditions, a load-optimised reference part was developed in the subsequent step. A finite element analysis (FEA) and a topology optimisation were carried out to determine the optimum material distribution in the reference part in terms of the load. The reference part was redesigned on the basis of these simulation results. This already enabled an initial lightweight design potential to be raised. In addition, the simulation results served as a preliminary stage for analysing the possible position and alignment of the continuous fibre reinforced inserts.

Before the design of the continuous fibre reinforced inserts could begin, it was necessary to evaluate the position in the injection-moulded part in which the continuous fibre reinforced inserts would be over-moulded and how much space they would take up in the geometry of the injection-moulded part. The shape of the continuous fibre reinforced inserts was restricted by the geometry of the injection-moulded part and should be aligned with the stress distribution in the part [13, 14]. For this reason, the design of the continuous fibre reinforced inserts was based on the results of the FEA calculation, in addition to the restrictions imposed by the space and the manufacturing constraints resulting from the additive manufacturing process. In the material model the anisotropy resulting from the injection moulding process also had to be taken into account. [15]

The next step was the virtual development of the hybrid reference part, which consists of two parts: the load-optimised injection-moulded part and the continuous fibre reinforced insert. In this stage of the development various virtual product development tools were used to design the hybrid reference part. An iterative process, combining FEA, topology optimisation and redesign, was run through to determine a load-optimal design for the hybrid reference part. The focus was on further weight savings by optimising the injection-moulded part and determining the geometry and optimal number of layers of the continuous fibre reinforced inserts. The strength of the hybrid reference part was simulatively validated using FEA For the simulation of the mechanical behaviour of the continuous fibre reinforced insert a non-linear anisotropic material model was used. (see Figure 3)

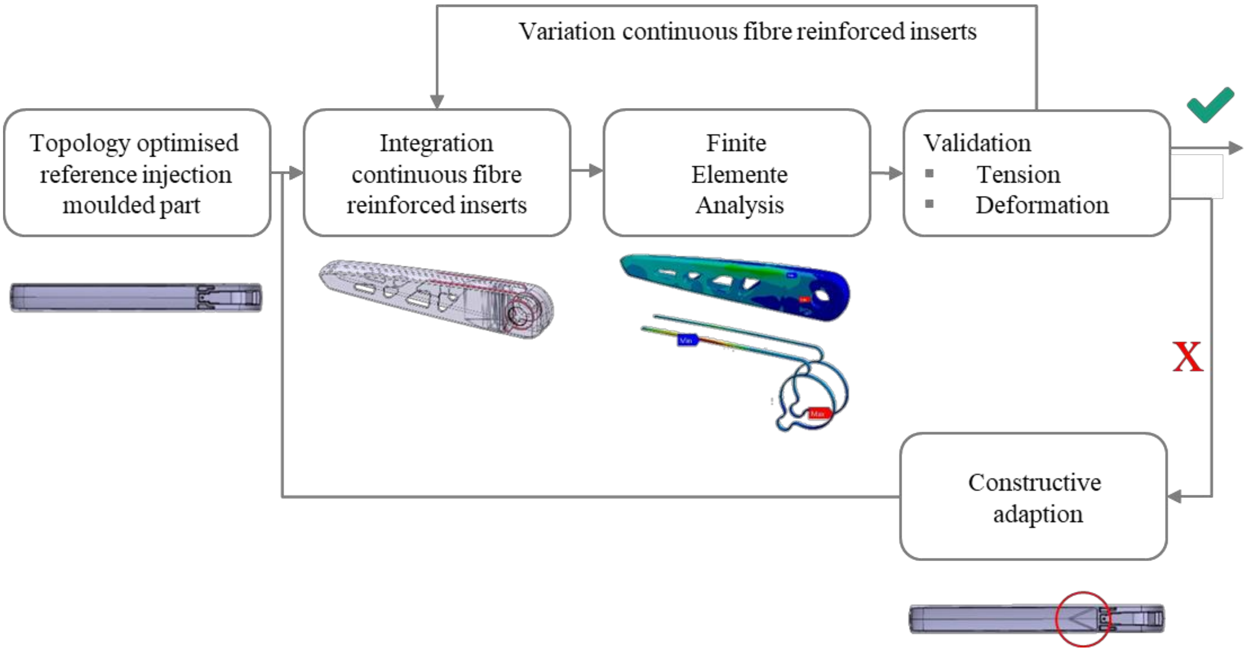


Figure 3: Procedure of virtual product development of the hybrid part

The optimum resulted from the cost aspect, whereby the size of the continuous fibre reinforced inserts was predominant, the insert costs were significantly higher than those of the injection-moulded part. Therefore, the optimal number of layers was primarily determined. The number of layers was considered optimal if the hybrid part reached the minimum safety factor and a further increase in the number of layers would not result in a significant improvement in strength or a reduction in the safety factor observed.

At the beginning, a topology optimisation of the reference part, which was normally manufactured by injection moulding, was carried out. The results of the topology optimisation were used as an indication of the areas where material and corresponding weight could be saved. Subsequently, the newly designed topology optimised hybrid part had to be validated by performing FEA calculations for the critical load cases. The boundary conditions were the same for all FEA calculations. The safety factor of the topology optimised part was used as the primary decision criterion for evaluating the strength of the hybrid part. Here a minimal safety factor had to be achieved so that the material did not fatigue under loading cases. Especially fibre fracture have to be prevented . For the further procedure for integrating the continuous fibre reinforced inserts, the determined load optimised structure served as the design basis of the hybrid part. The results of this FEA calculation were used to locate the optimal positions for the continuous fibre reinforced inserts.

Due to the load distribution and the formation of weld lines in the part, the area of the armrest's bearing surface and the area of the assembly suspension on which the armrest was mounted to the car seat were identified as critical. Therefore, the continuous fibre reinforced inserts were used in these areas to further reinforce the part. The main objective was to ensure that the load flow when the armrest was under stress was primarily directed into the continuous fibre reinforced inserts and thus into the fibres, which leaded to the load relief of the topology optimised structure. The next step was to determine the design of the continuous fibre reinforced inserts. The design was limited on the existing volume of the topology optimised injection-moulded part and the manufacturing restrictions of the continuous fibre reinforced additive manufacturing process (layer height 0.27 mm, fibre width: 1.1 mm). In the first instance, the optimum number of layers of the continuous fibre reinforced inserts were determined, starting with a minimum number of layers of two. Subsequently, the structure of the hybrid part was validated by means of renewed FEA calculations. This involved checking whether the topology optimised structure together with the continuous fibre reinforced inserts (hybrid part) was sufficient for the existing loads. In particular, the focus here was on the continuous fibre reinforced inserts, where it was important to avoid fibre breaks so that the load flow continued to be directed into the fibres. If the FEA calculation of the topology optimised hybrid part results insufficient safety factors the continuous fibre reinforced inserts were primarily redesigned by increasing the number of layers. If necessary, many iteration loops had to be performed until the optimum number of layers was achieved and the requirements of the topology optimised hybrid part could be observed simultaneously.

However, if the threshold of the safety factors and the total deformation could not be exceeded or not reached and if there was no more space to increase the number of layers of the continuous fibre reinforced inserts, a constructive adaptation of the topology optimised part had to be carried out. This meant that too much material was removed in the first step of the topology optimisation and the continuous fibre reinforced inserts could not compensate loss of strength due to the material reduction.

Therefore, an improvement of the stability of the topology optimised hybrid part had to take place by a constructive modification of the topology optimised structure. At this point, the iteration loop for determining the optimal number of layers and observation with the requirements for the topology optimised hybrid part started all over again. If, however, the optimal number of layers and the requirements for the part were achieved, it might also be possible to save material and weight by reducing the material additionally. The weight saving was to be chosen as the termination criterion of the iteration loop. As soon as no significant weight advantage could be achieved by the constructive modification of the topology optimised structure, the iteration loop was terminated and the continuous fibre reinforced insert did not result in a significant improvement of the strength. The procedure was characterised by numerous iteration loops until a result could be achieved. These iteration loops were carried out until the constructive redesign of the topology optimised structure showed no weight advantage compared to the reference part.

Following the virtual validation of the topology optimised hybrid part, the physical validation procedure of the process chain for hybrid parts was carried out. After the continuous fibre reinforced inserts had been additively manufactured, they were inserted into specially provided and appropriately designed cavities in the injection mould. The inserts were manufactured from a continuous fibre-reinforced polymer filament, composed of 60 wt.-% 3K fibre filament count tows and 40 wt.-% Polyamide 12 polymer. Then the continuous fibre reinforced inserts were over-moulded with polymer (Polyamide 66 with 30 wt.-% glass fibre content). Otherwise, the process did not differ from the conventional injection moulding cycle. It was essential for a good injection moulding result that the continuous fibre reinforced inserts were suitably fixed, because otherwise the complete encapsulation by the melt was jeopardised and warpage could occur [1].

Furthermore, additional criteria had to be observed when the polymer melt flowing around the inserts. Ideally, the position of the continuous fibre reinforced inserts should be chosen advantageously in relation to the weld line positions in order to avoid air inclusions. Therefore, the flow of the melt around the continuous fibre reinforced inserts had to be optimised. In addition, the reorientation of the long glass fibre of the injection moulding material had to be taken into account, which occurred when flowing around the inserts by the polymer melt. The long glass fibres aligned along the continuous fibre reinforced inserts and enveloped these. The flow direction of the melt also helped pre stretching the continuous fibres in the load direction [4].

In the last step of the development several selected prototypes of the hybrid reference parts were manufactured, which have different modifications of the continuous fibre reinforced inserts (path design of the continuous fibre or the number of layers of the inserts). In this process, continuous fibre reinforced inserts were additively manufactured from a composite filament using the continuous fibre reinforced MEX process AFT. The continuous fibre reinforced inserts were placed in an injection moulding tool and over-moulded with long glass fibre reinforced thermoplastic. By attaching polymer distance keepers to the continuous fibre reinforced inserts, it was possible to ensure that they hold their position in the injection mould. Finally, the hybrid parts were validated by means of a test set-up that reproduced the real load situations in the mounted state. This was used to validate the results from the simulation and to qualify the development process for series production.

4. Results

The technical analysis of the hybrid part compared to the reference part showed a 19.5% reduction in weight in the primary objective of weight saving. In the first critical load case (load from above), a 38.1% lower deformation was achieved (see figure 4). The specified maximum load and deformation limits resulting from the application were adhered too. In addition, failure in the area of the weld line could be avoided due to the continuous fibre reinforced inserts.

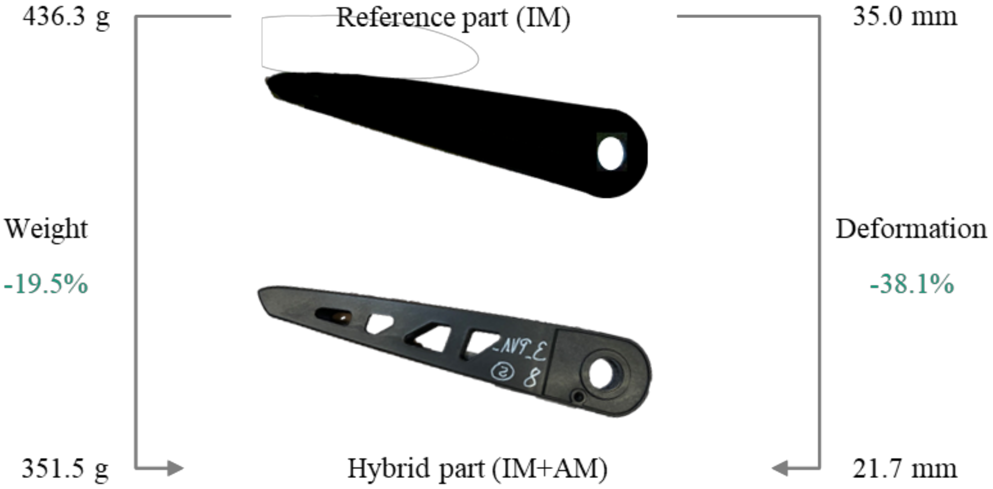


Figure 4: Comparison of reference part (IM) and hybrid part (IM + AM)

During the simulations, it was observed that an increase in the number of layers of the continuous fibre reinforced insert up to an optimum has a positive influence on the increase in the strength of the hybrid part. Further layer number increases beyond the optimum result in a reduction of the strength. This can be explained by the interlaminar interaction of the individual layers, whereby delamination takes place and the detachment of the individual layers leads to an overall reduced reinforcing effect of the continuous fibre reinforced inserts. The delamination beyond the optimum layer number of the inserts occurs because of the reduced layer bonding during long manufacturing time. The long manufacturing time results in inhomogeneous cooling conditions, which have a negative impact on layer bonding. In addition, the weak point of the weld lines could be compensated by the reinforcing effect of the continuous fibre reinforced inserts. A shift of the fracture points away from the critical area of the assembly fixture to the front area of the armrest was observed. The economic analysis of the hybrid part showed that the costs were still too high for the application in the automotive industry. The value of 5 €/kg additional costs per saved kilogram in the prototypes, which is usual for the automotive industry, was exceeded. However, it is an attractive solution for special solutions in vehicles or for applications in medical technology or the aviation industry.

5. Summary and Outlook

To answer the research question, which focuses on achieving lightweight construction goals in a long glass fibre reinforced injection-moulded part, the concept of continuous fibre reinforcement was developed. The lack of processing possibilities for continuous fibres within the injection moulding process made it necessary to add another manufacturing process to integrate the continuous fibres (process chain for hybrid parts). The continuous fibres were implemented as semi-finished products according to the principles of insert technology. For the production of these continuous fibre reinforced inserts, the continuous fibre reinforced additive manufacturing was selected. An iterative procedure for the virtual product development of a hybrid part could be worked out. Especially the interaction of different tools of virtual product development from design, FEA validation and simulation has an innovative character. Furthermore, it should be critically noted that the anisotropic elastic constants and strengths of the parts and materials for the FEM and simulation were determined by semi-empirical equations, since no reliable material parameters were available from the manufacturers or in the literature. More realistic simulation results can be expected if the anisotropic elastic constants and strengths are determined experimentally by means of tensile tests. In addition, the quality of the simulation results can be increased if there is a better interface between injection moulding and FEM simulation. In this context, further developments of the software are to be made with regard to the consideration of continuous fibre reinforced laminates in the injection moulding simulation according to the principles of insert technology.

Nevertheless, the validation of the hybrid parts showed that the simulation results correspond to the real load situation in terms of strength. In the hybrid parts, a further increase in the mechanical part properties could be achieved by optimising the continuous fibre reinforced inserts. The technical analysis of the hybrid part compared to the reference part showed a 19.5% reduction in weight in the primary objective of weight saving. In the first critical load case (load from above), a 38.1% lower deformation was achieved. An additional adaptation of the injection moulding tool is not necessary for this. In particular, the optimal increase in the number of layers to stabilise the continuous fibre reinforced inserts in the injection moulding tool. Likewise, the strength of the hybrid part can be further increased by compressing the continuous fibre reinforced inserts (improved layer adhesion of the individual continuous fibre layers).

The bonding of the inserts with the injection-moulded matrix of the hybrid part creates a material, form and friction bond without the need of a post-processing step. Whereas with metal inserts only a friction or form bond can be created and for material bonding a post processing step is necessary. This is favoured by a low heat distortion temperature of the polymer matrix of the continuous fibre reinforced inserts, so that the polymer matrix of the continuous fibre reinforced inserts is also melting during over moulding with the long glass fibre reinforced thermoplastic. This results in an atomic bond between the two polymer matrices when the part solidifies.

Considering the economic circumstances of the application in automotive industry the part costs to achieve this weight saving are considered too high, if the acceptable additional costs per weight saved in the automotive industry (5 €/kg) serve as a reference. Therefore, the application in aviation industry in order to substitute metallic parts through hybrid parts can be considered.

Furthermore for more applications the superior thermal and electrical properties of the continuous fibre reinforced inserts can be utilised. Besides the enhancement of mechanical properties and the weight saving potential the continuous fibre reinforced inserts can reduce thermal expansion and increase the thermal conductivity for the application in the optics industry. The high electrical conductivity of the continuous fibre reinforced inserts can also be used to integrate sensors and selective conducting paths in injection moulding parts.

Acknowledgements

This research was supported by the European Fund for Regional Development (EFRE) and the Oberfrankenstiftung.

References

[1] Koch, T.; Schürmann, H.; 2006; Spritzgussbauteile lokal verstärken. In: Kunststoffe

[2] Suzuki, T.; Fukushige, S. and Tsunori, M.; 2020; Load path visualization and fiber trajectory optimization for additive manufacturing of composites, Additive Manufacturing, vol. 31, p. 100942,

[3] Dutra, T. A.; Ferreira, R. T. L.; Resende, H. B. and Guimarães, A.; 2019; Mechanical characterization and asymptotic homogenization of 3D-printed continuous carbon fiber-reinforced thermoplastic, J Braz. Soc. Mech. Sci. Eng., vol. 41, no. 3, p. 133

[4] Mohammadizadeh, M.; Imeri, A.; Fidan, I. and Elkelany, M.; 2019; 3D printed fiber reinforced polymer composites - Structural analysis, Composites Part B: Engineering, vol. 175, p. 107112

[5] Dickson, A. N.; Barry, J. N.; McDonnell, K. A. and Dowling, D. P.; 2017; "Fabrication of continuous carbon, glass and Kevlar fibre reinforced polymer composites using additive manufacturing," Additive Manufacturing, vol. 16, pp. 146–152

[6] Czasny, M.; Goerke, O.; Kaba, O.; Koerber, S.; Schmidt, F. and Gurlo, A.; 2019; Influence of Composition on Mechanical Properties of Additively Manufactured Composites Reinforced with Endless Carbon Fibers, KEM, vol. 809, pp. 335–340,

[7] Tian, X.; Liu, T.; Yang, C.; Wang, Q. and Li, D.; 2016; Interface and performance of 3D printed continuous carbon fiber reinforced PLA composites, Composites Part A: Applied Science and Manufacturing, vol. 88, pp. 198–205

[8] Domm, M.; Schlimbach, J. and Mitschang, P.; 2019; Optimizing mechanical properties of additively manufactured FRPC, 21st International Conference on Composite Materials, Xi'an, p. 12

[9] Chacóna, J.M.; Caminerob, M.A.; Núñezb, P.J.; García-Plazab, E.; García-Morenob, I.; Revertea, J.M.; 2019; Additive manufacturing of continuous fibre reinforced thermoplastic composites using fused deposition modelling: Effect of process parameters on mechanical properties", Composites Science and Technology, Vol. 181

[10] Wang, X.; Jiang, M.; Zhou, Z.; Gou, J.; Hui, D.; 2017; 3D printing of polymer matrix: a review and prospective, Composites: Part B 110, p. 442–458.

[11] Pezold, D.; Rosnitschek, T.; Kleuderlein, A.; Döpper, F.; Alber-Laukant, B.; 2021; Evaluation of Technologies for the Fabrication of Continuous Fibre Reinforced Thermoplastic Parts by Fused Layer Modeling. in: Technologies for economic and functional lightweight design, Springer

[12] 9TLabs AG: Carbon Composite Material. https://www.9tlabs.com/technology/material Accessed on 04.02.2022

[13] Menges, G.; Michaeli, W.; Mohren, P.; 2007; Spritzgießwerkzeuge. München: Carl Hanser Verlag

[14] Rohde-Tibitanzl, M.; 2015; Direct Processing of Long Fiber Reinforced Thermoplastic Composites and Their Mechanical Behavior under Static and Dynamic Load. 1. Auflage. München: Carl Hanser Verlag GmbH & Co. KG

[15] Schürmann, H.; 2007; Konstruieren mit Faser-Kunststoff-Verbunden. 2. Auflage. Heidelberg: Springer Verlag

Biography

Johann Schorzmann (*1991) is a research associate at the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth since 2020. He has managed and conducted a number of research and industrial projects in the context of additive manufacturing.

Daniel Pezold (*1988) started working at the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth in 2017. In his current position within the group of additive manufacturing, he has the project lead of "Anwendungszentrum 3D-Druck Oberfranken". He is specialized within endless fibre reinforced polymers and High Speed Sintering.

Jan Kemnitzer (*1989) is working at Fraunhofer Institute for Manufacturing Technology and Automation IPA in Bayreuth. He is group leader of the division additive manufacturing and is specialised on the powder bed fusion of polymer (PBF-P) processes.

Alexander Kalusche (*1962) founded and is leading the acad group GmbH in Heilsbronn. Acad group is specialised in rapid fibre reinforced polymer injection molding and is designing polymer parts for the automotive and medical sector.

Frank Döpper (*1970) has been head of the Chair of Manufacturing and Remanufacturing Technology at the University of Bayreuth and the Fraunhofer Institute for Manufacturing Technology and Automation IPA in Bayreuth since 2017.

3rd Conference on Production Systems and Logistics

Spare Parts Demand Forecasting in Maintenance, Repair & Overhaul

Torben Lucht[1], Volodymyr Alieksieiev[1], Tim Kämpfer[1], Peter Nyhuis[1]
[1] *Leibniz University Hannover, Institute of Production Systems and Logistics (IFA), Germany*

Abstract

Despite a high degree of uncertainty about the scope of future orders and the corresponding capacity and material demands, Maintenance, Repair & Overhaul (MRO) service providers face high expectations regarding due date reliability by their customers. To meet these requirements while at the same time keeping delivery times short, the availability of the required spare parts or pool parts is an essential success factor. As these cannot be kept in stock in large quantities due to their high monetary value, reliable spare parts demand forecasts are of vital importance for the profitability of MRO service providers. As a result of a high degree of information uncertainty and the mostly lumpy demand patterns, conventional time-based and statistical methods do not show sufficient forecasting quality for application in the MRO industry. Data-based approaches incorporating machine learning methods offer promising capabilities to achieve improved predictive accuracy but still need to be adequately linked to production planning and control to realize their full potential. This paper first analyses potential approaches to spare parts demand forecasting in the MRO industry, focusing on forecast accuracy and potential for integration into material and production planning. Based on this, a classification of demand forecasting approaches is presented and an approach for order-based material demand forecasting with two-step feature selection is proposed. Finally, the presented approach is applied on a real dataset provided by an MRO service provider.

Keywords

MRO; spare parts demand; forecasting; Machine Learning; Artificial Neural Networks.

1. Introduction

Maintenance, Repair and Overhaul (MRO) of complex capital goods, such as aero engines or wind turbines, is also known as “regeneration” [1]. This process comprises the disassembly, inspection, repair, reassembly, and test (quality control) of mostly high value products [2]. In addition to this, there are up to two pooling stages in the regeneration supply chain (see Figure 1) to provide repairable or serviceable spare parts to their downstream processes and by this improve robustness against disturbances or material shortage along the regeneration process. [3]. These pools are filled either from the respective upstream processes or via the procurement of new or used parts. The availability of the pool parts and the precision of the corresponding demand forecast thus have a significant influence on the punctuality of the material supply for the reassembly and the achievable adherence to delivery dates of the MRO service provider to its customers [4]. In turn, the on-time delivery by MRO service providers is complicated with the high degree of uncertainty about the future work scope at the beginning of the regeneration process. Due to the complexity of goods to be repaired, it is not possible until the end of the inspection to recognize all existing damages and thus to plan repair operations and forecast the material demand. Furthermore, it is uncertain, whether a component can be repaired or has to be replaced (e.g. due to heavy damage) [1].

DOI: https://doi.org/10.15488/12179

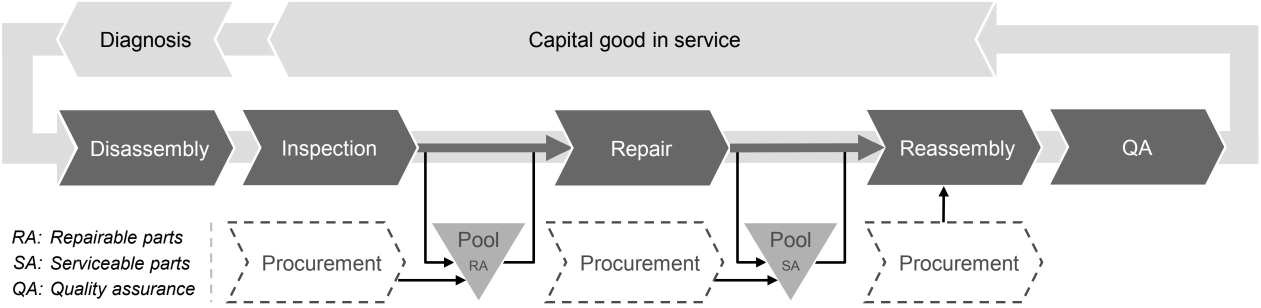


Figure 1 – Universal supply chain structure for the regeneration of aero engines [5]

As the spare parts cannot be kept in stock in large quantities due to their high monetary value, reliable forecasting is a crucial factor to ensure the profitability of the MRO service provider. Because of the lumpy patterns of spare parts demand, which will be described in the next section, traditional time-series and statistical forecasting methods do not provide sufficient forecasting quality for application in the MRO industry [6]. However, today more and more condition data, e.g. oil pressure or temperatures are measured during operation, which can be indicators regarding the wear of components [7]. Besides these quantitative parameters, also qualitative parameters, such as region, climate, maintenance politics of aircraft operator or owner have to be considered while forecasting material demand. This is possible e.g. using Machine Learning (ML)-based methods, which thus are the focus of this paper. Based on a brief introduction to spare parts demand classification a brief analysis of characteristics of spare parts demand in the MRO industry and potential methods for spare parts demand forecasting is performed in section 3. Based on this, section 4 presents a hybrid approach to spare parts demand forecasting and outlines the first prediction results obtained. Finally, conclusions are given in section 5.

2. Spare parts demand in the MRO industry

Spare parts demand can be categorized, using periodicity (inter-demand intervals) and quantity variation. Typical demand structures are smooth, erratic, intermittent and lumpy demand (see Table 1) [8,9].

Table 1 – Demand categorization according to [8], [9]

Demand Type	Inter-demand intervals	Quantity variation
Smooth	Low	Low
Erratic		High
Intermittent	High	Low
Lumpy		High

Smooth and erratic demand patterns can be distinguished according to quantity variation, which is relatively low in the case of smooth demand patterns and relatively high in case of erratic demand. Periods between demand occurrence are small in both cases. Intermittent and lumpy demand is characterized by the mostly random appearance of demand and many periods of zero demand. Furthermore, lumpy demand, in comparison to intermittent demand, shows high variance in spare parts quantity [11,10]. Cut-off values regarding the separation of these demand patterns are proposed in [10]. Considering complex capital goods like aircraft about 80% of the demand for repair, and corresponding material demand comes up unplanned [12]. Due to this and corresponding uncertainties regarding damage pattern, work scope and spare parts demand of unplanned MRO-activities can mostly be categorized as intermittent or lumpy (cf. [6] for sources of intermittency and lumpiness for aircraft spare parts). Hence, different forecasting methods and potential fields of application in forecasting of intermittent or lumpy demand are analyzed in the next section.

3. Literature review: Forecasting of material demand

Methods for demand forecasting methods overall can be grouped in deterministic, stochastic demand assessment and subjective estimation methods [13]. [14] categorizes forecasting approaches depending on the influencing variables in causal, lifecycle, time series and consumption analysis. A differentiation between qualitative and quantitative approaches is used in [15], whereby the quantitative methods are subdivided in uni- and multivariate methods. [16] uses a similar structure but subdivides quantitative methods in time-series and causal forecasts. An alternative classification is presented in [17] that distinguishes between past-based and future-based methods, each divided into qualitative and quantitative methods. These are further differentiated in methods for forecasting of time and quantity of material demand by [18]. These approaches to classification of material demand forecasting form the basis for the classification scheme (see Figure 2) that is presented in the following sections.

3.1 Deterministic approaches

Deterministic demand forecasting methods are methods by which material demand is determined solely based on an existing independent primary demand [13]. These methods comprise analytical and synthetic approaches [13]. Analytical methods rely on the bills of material of the finished product. Based on them, the demand on finished product (primary demand) is disassembled in demand for subassemblies and components [13]. Synthetical methods to forecasting make use of parts usage lists as a forecast basis and are suitable especially for long-term planning [13]. Another deterministic approach is e.g. consumption analysis. This method is based on maintenance measures planning [14]. Due to their inability to consider uncertainties and thus unplanned material demand in the regeneration process, deterministic approaches are only suitable for spare parts provision during planned regeneration events (e.g. mandatory replacements of components). For intermittent and lumpy demands, which are in the focus of this paper, stochastic methods are commonly used [19].

3.2 Stochastic approaches

Stochastic demand forecasting can be defined as "mathematical-statistical methods, in which past consumption values are used to infer future demand" [13]. These methods can again be grouped into quantitative and qualitative. Quantitative stochastic methods include univariate and multivariate approaches [15] that are presented separately in the following sections.

3.2.1 Quantitative univariate approaches

Univariate approaches are those based on consideration of only one independent variable and include e.g. time-series and life-cycle analysis. Time-series methods are methods by which the forecasting for a future time horizon is made based on a demand history from the past. Among others, the approaches based on well-known statistical methods, such as exponential smoothing or moving average, are to be emphasized. Statistical methods for forecasting intermittent and lumpy demands were first studied by CROSTON [20]. In his work, he found that exponential smoothing does not provide sufficient forecast quality to forecast intermittent demand and proposed his method, based on exponential smoothing, in which demand rate and time intervals between its occurrence are analyzed and forecasted separately [20], [21]. [10] and [22] identified a bias in CROSTON's method and introduced an additional correction factor to avoid this bias. Further statistical methods for predicting intermittent and lumpy demand are also presented and discussed in [23], [24], [25] and [26].

Life-cycle analytical methods for demand forecasting are based on an "estimation of the time until failure of the corresponding component" [14]. These methods are based on failure rates or, in other words, the probability of a failure as a function of its lifetime [15]. Practical studies on these methods are presented e.g. in [27] and [28].

Time-series and life-cycle analytical methods are easy to use and require a relatively small amount of input data. Nevertheless, the increasing number of influencing factors that MRO service providers are provided with, e.g., from condition monitoring systems, cannot be taken into account completely with the help of these approaches, which leaves potential for improvements of the forecast quality unused.

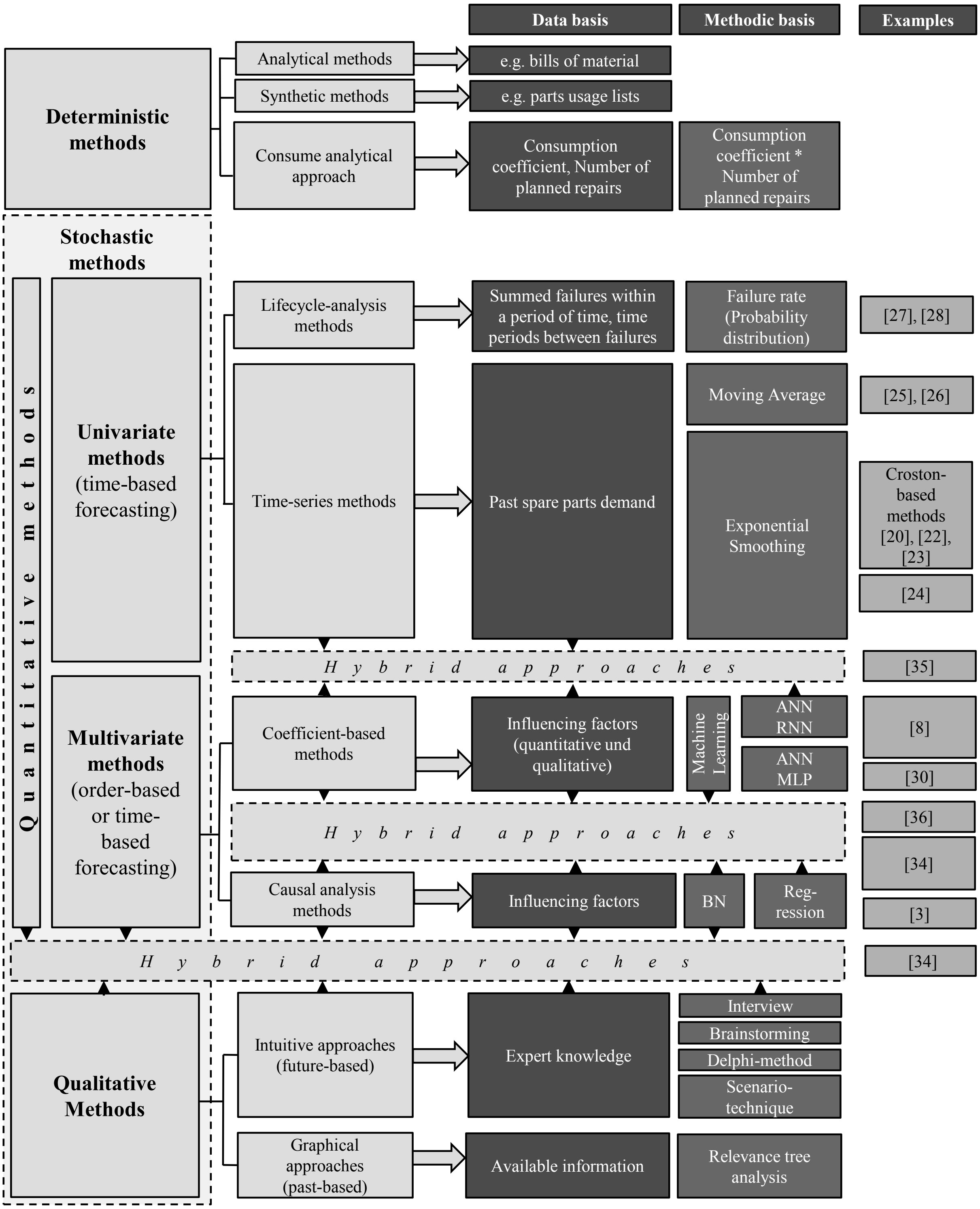


Figure 2 – Classification of demand forecasting methods (based on [13], [14], [15], [16], [17], [18])

3.2.2 Quantitative multivariate approaches

As forecasting material demand is usually dependent on more than one variable, multivariate forecasting methods are gaining more and more importance over the recent years. These methods include coefficient-based and causal analysis methods [15]. These approaches typically apply data from the use phase, for example using condition monitoring systems, or the maintenance phase of the goods (cf. [7]). Coefficient-based methods consider several influencing factors (quantitative and qualitative) to determine a wear coefficient (cf. [15] for definition). These methods include, for example, ML-based methods, such as Artificial Neural Networks (ANN) that represent simplified representations of the biological neural network (cf. [29] for the definition of ANN). They consist of several information processing units ("neurons") that contain mathematical functions and are interconnected. The signals entering a neuron are weighted and converted into the output signals using an activation function. To do so, the ANN is trained based on a training data set, e.g. to achieve desired prediction results. Lumpy demand forecasting using a multilayer perceptron (MLP) type of ANN is explored and analyzed in [8], [30], and [31]. The analysis of 60 contributions related to ANN-based intermittent demand forecasting in [8] reveals, that MLP-based methods provide the best forecasting performance compared to other types of ANN. Above mentioned research also proves that the forecasting accuracy of MLP outperforms that of time series analytical methods. Other ANN-based methods for forecasting material demand are investigated in [32] (e.g. Recurrent Neural Networks (RNN)), that also show good results in the forecasting of non-stationary demand in the field of aircraft spare parts management. Through good forecasting performance, big input-data requirements as well as poor traceability auf causal relationships can be highlighted as disadvantages of ANN-based methods. These can be identified using causal analysis forecasting methods [15]. One of the most common causal forecasting methods are Bayesian networks (BN). BN are a set of variables (nodes) and directed edges between them, that form a directed acyclic graph (DAG). Edges of this graph represent potentially causal dependencies between the nodes [3], [33]. First applications of different types of BN (expert-initiated BN, data-based BN, and hybrid ML-based BN, which combines the first two approaches) for forecasting lumpy spare parts demand are performed in [34]. Here, the hybrid BN outperforms the expert-initiated BN and the data-based BN as well as logistic regression in terms of prediction accuracy [34]. First applications of BN in regeneration logistics can be found in [3]. In this context, they are used to determine the probability with which regeneration orders are required for a component or an assembly. For this purpose, the product structure of the regeneration good is represented in form of a BN, in which the assemblies and components are mapped as its nodes. The edges are derived from the product structure and existing influencing factors. For the determination of the initializing probability distribution, existing service data from the past is provided as the basis for the BN. In both [3] and [34] good causality determination performance of BN is reported. In contrast to time-series based methods, quantitative multivariate forecasting can be used for order-specific forecasting to predict the demand for a certain regeneration order, e.g. based on conditional or operating data of a certain regeneration good. However, this only allows the total demand per order to be predicted and does not include information about the specific time this demand occurs. This is illustrated graphically in Figure 3 using a fictitious demand time-series.

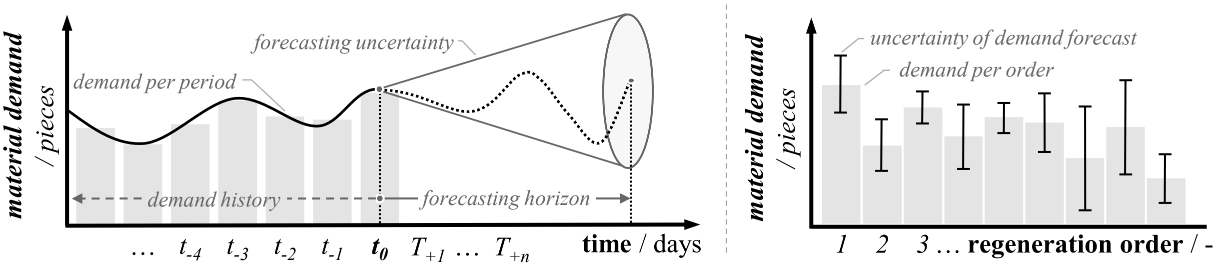


Figure 3 – Types of material demand forecasting

3.2.3 Qualitative approaches

Qualitative approaches are methods based on expert estimates or the analysis of existing information (without causality determination) about the forecasting asset. These can be subdivided in past-based and future-based qualitative forecasting methods. Methods based on past data include, for example, relevance tree analysis. The future-based methods include, among others, questioning, brainstorming, Delphi method, and scenario technique. [17]. Although qualitative methods are widely used for spare parts demand estimation in the MRO sector due to their simplicity, they are still strongly dependent on individual, subjective estimations and thus can neither be proven by data, nor can they be reproduced or even automated. Due to high financial risks, the high variability in demand as well as the complexity of the goods, the quality of the forecasts is often insufficient, which is why they are not the focus of this paper.

3.3 Hybrid approaches to material demand forecasting

Hybrid forecasting approaches combine different forecasting methods to improve forecast accuracy. In [35] hybrid approach for intermittent demand forecasting in the semiconductor supply chain is proposed, which combines RNN-based and time-series-based methods. In this study, the presented method outperforms time-series and RNN-based forecasting methods in terms of demand prediction accuracy. In [36] a hybrid approach for material demand forecasting dedicated to the mining industry is proposed. It combines regression modeling and ANN-based method and which also shows better forecasting performance compared with time-series and ANN-based methods as standalone approaches.

The overview of relevant literature has shown, that advanced ANN MLP-based approaches outperform conventional statistics methods in forecasting accuracy. Hence, in the following section an ANN-based order-specific approach dedicated to the MRO industry is presented. This order-specific forecast could afterwards potentially be distributed over the demand time periods, which could be a topic of further research.

4. Overview of ANN MLP-based approach for material demand forecasting

As mentioned in section 3, ML-based and, especially, ANN MLP-based methods provide better forecasting performance in comparison to the time-series methods. In this section, hence, an approach for systematic application of ANN for order-specific material demand forecasting in the MRO industry is presented. First the approach functionality and general process is presented in section 4.1. Afterwards its software-based implementation based on real dataset provided by MRO service provider is presented in section 4.2.

4.1 Overview of approach functionality

The performance of ANN MLP-based methods can be significantly improved by the selection of relevant input-features (cf. [38,37]). The approach presented in this section (see Figure 4) is focused on sufficient data preparation and feature selection for ANN MLP-based order-specific demand forecasting for the MRO industry. It combines qualitative and causal analysis methods into a two-step process to select relevant features and, by this, increase forecasting accuracy. The structure of the approach is based on typical structure of data analytics project, presented e.g. in [39]. Consequently, the first step of the approach is data preparation based on typical datasets available to MRO service providers. This usually comprises condition parameters, contractual information, customer related data and data from previous regenerations of similar or the same product. To apply ANN-processing this data needs to be prepared accordingly (e.g. through normalization). Afterwards the data irrelevant to the subject area needs to be excluded – usually in corporation with a subject area expert (e.g. internal customer numbers). This may help to decrease computational time and costs for the next FS-step (see Figure 4). After this assessment and basic filtering of irrelevant features, systematic techniques for feature selection have to be applied to avoid redundancy [37], which can not be identified during expert evaluation, as well as to enhance the understandability and to minimize the effort of further data processing [38].

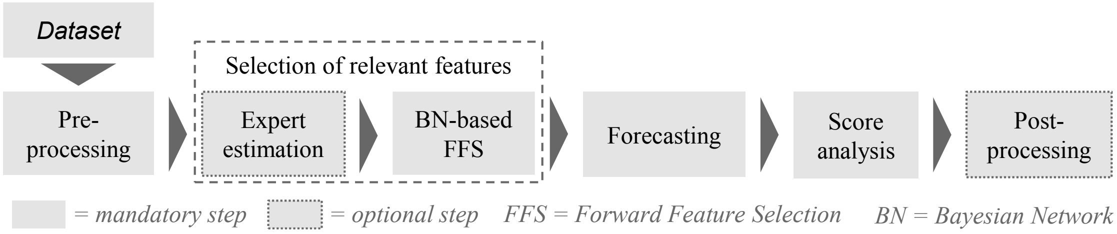


Figure 4 – ANN MLP-based approach for material demand forecasting in the MRO industry

This paper focuses on Forward Feature Selection (FFS) only as one of the most popular feature selection methods. This represents an iterative approach, that progressively adds features that improve the model's forecasting accuracy the most until no additional accuracy can be gained [40]. Due to the increasing number of features available in regeneration this needs to be supported systematical. To do so, BN are chosen as a model learner, due to their good performance in the identification of interdependencies as reported in [3] and [34]. After relevant features have been selected, forecasting can be performed and analyzed using statistical failure rates. This assessment allows for a preliminary evaluation of forecasting results. If the applied forecasting method did require the normalization of data during pre-processing, data has to be denormalized to obtain forecast values usable in practice.

4.2 Software-based application of ANN-based order-specific demand forecasting

For validation of the proposed approach functionality, it was applied to a real data set, provided by an MRO service provider. The data provided comprises more than 600 datasets with 22 qualitative and quantitative parameters each. The data preparation and forecasting method were implemented using the open-source visual-programming tool *KNIME Analytics Platform v4.5*. Using above described two-step-FS the following features were selected: cycles since new and last regeneration and (partially) product owner, region of operation, regeneration project type. To analyze the forecasting accuracy the results obtained by forecasting with one-step (only expert estimation) feature selection is compared with results obtained using the presented two-step (expert estimation and FFS) feature selection based on typical statistical measures: Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) (see Table 2). In this comparison the normalized values are used for better understanding of the range of the forecasted value. The training algorithm was repeated ten times to determine the achievable range of forecasting accuracy.

Table 2 – Comparison of forecasting accuracy of ANN MLP with one-step and two-step of feature selection

Error	MAE	MSE	RMSE
ANN MLP (1 St.)	0,292…0,300	0,115…0,135	0,339…0,367
ANN MLP (2 St.)	0,288…0,299	0,112…0,120	0,340…0,346

The comparison confirms that ANN MLP-based approach with two stages of feature selection outperforms the similar order specific approach with only one step of feature selection (expert evaluation) in forecasting accuracy. It needs to be mentioned, that in this example only required demand for serviceable components were forecasted, as there was no information on capacity demands per regeneration order, which have to be included for demand forecasting and demand-oriented inventory dimensioning of repairable spare parts. For comparison with conventional time series-based approaches, it also has to be taken into account that the prediction results obtained with the ANN-MLP approach so far only forecast order-specific demands without their demand timing. Consequently, it requires a scheduling of the demands based on the probability of occurrence of the regeneration events as well as the delay in demand based on the date of occurrence of the regeneration events. A potential approach to this estimation is described in [41] that uses a hybrid approach of data mining and logistics models to predict throughput times of regeneration orders. As mentioned in section 4, this coupling should be focused next to allow for an application in the MRO industry.

5. Summary and outlook

Despite various research regarding the prediction of mostly intermittent or lumpy spare parts demand in the MRO-industry service providers still lack suitable and applicable approaches to spare parts demand forecasting using available quantitative and qualitative information. In this paper, different methods for material demand forecasting are analyzed, compared and systematically structured. Here it needs to be differentiated between time-based and order-based forecasting. The literature review has shown, that ML- and especially ANN-based forecasting methods significantly outperform conventional time-series methods in terms of forecasting non-stationary demand. Taking into account MLP as the best performing approach among other ANN-based methods, a systemic approach for application of ANN MLP to forecast material demand in the MRO industry was proposed afterwards. Further this approach was applied to a real dataset provided by an MRO service provider for the prediction of required quantity of serviceable components with two stages of feature selection (expert estimation and FFS). Its performance was compared with the similar approach, using one-step feature selection (expert estimation) only, afterwards. This comparison has shown, that using two-stage feature selection with FFS technique, based on a BN learner, better forecasting accuracy can be achieved. Further research needs to be dedicated to the hybridization of time-based and order-based forecasting approaches with the purpose of distributing precise ANN-based demand forecasts over time periods. In this context, the material demand forecast must also be extended to include the expected demand for repair, so that inventories of repairable components can also be systematically taken into account for the purpose of meeting the total material demand. An additional direction of research is the comparison of alternative feature selection methods and different selection model learners to further improve forecast accuracy.

Acknowledgements

Funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) ”SFB 871/3” - 119193472.

References

[1] Eickemeyer, S.C., 2014. Kapazitätsplanung und -abstimmung für die Regeneration komplexer Investitionsgüter. Dissertation, Hannover.

[2] Guide, V.R., Kraus, M.E., Srivastava, R., 1997. Scheduling policies for remanufacturing. International Journal of Production Economics 48 (2), 187–204.

[3] Berkholz, D.A., 2012. Grundmodell zur Kapazitäts- und Belastungsabstimmung eines Arbeitssystems in der Regeneration. Dissertation, Hannover.

[4] Hermeier, B., Platzköster, C., 2006. Ergebnisse der ersten bundesweiten FOM-Marktstudie „Industrie-Dienstleistungen“, Essen.

[5] Lucht, T., Kämpfer, T., Nyhuis, P., 2019. Characterization of supply chains in the regeneration of complex capital goods, in: Dimitrov, D., Hagedorn-Hansen, D., Leipzig, K. von (Eds.), International Conference on Competitive Manufacturing (COMA 19) proceedings. 30 January 2019-1 February 2019, Stellenbosch, South Africa. Department of Industrial Engineering Stellenbosch University, Stellenbosch, South Africa, pp. 444–449.

[6] Ghobbar, A., Friend, C., 2002. Sources of intermittent demand for aircraft spare parts within airline operations. Journal of Air Transport Management 8 (4), 221–231.

[7] Andersson, J., Jonsson, P., 2018. Big data in spare parts supply chains. IJPDLM 48 (5), 524–544.

[8] Amirkolaii, K.N., Baboli, A., Shahzad, M.K., Tonadre, R., 2017. Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by using Artificial Intelligence (AI). IFAC-PapersOnLine 50 (1), 15221–15226.

[9] Eaves, A.H.C., Kingsman, B.G., 2004. Forecasting for the ordering and stock-holding of spare parts. Journal of the Operational Research Society 55 (4), 431–437.

[10] Syntetos, A.A., Boylan, J.E., Croston, J.D., 2005. On the categorization of demand patterns. Journal of the Operational Research Society 56 (5), 495–503.

[11] Syntetos, A.A., 2001. Forecasting of intermittent demand. Ph.D. Thesis.

[12] Mensen, H., 2013. Handbuch der Luftfahrt. Springer, Berlin, Heidelberg.

[13] Hartmann, H., 2002. Materialwirtschaft: Organisation, Planung, Durchführung, Kontrolle, 8., überarb. und erw. Aufl. ed. Dt. Betriebswirte-Verl., Gernsbach, 740 pp.

[14] Markiewicz, M., 1988. Bedarfsermittlung für Ersatzteile, in: Besters, H., Colbe, W.B. von, Engelhardt, W., Jaeger, A., Laßmann, G., Maßberg, W., Schwark, E., Wartmann, R., Markiewicz, M. (Eds.), Ersatzteildisposition im Maschinenbau, vol. 32. Gabler Verlag, Wiesbaden, pp. 35–67.

[15] Loukmidis, G., Luczak, H., 2006. Lebenszyklusorientierte Planungsstrategien für den Ersatzteilbedarf, in: Barkawi, K., Baader, A., Montanus, S. (Eds.), Erfolgreich mit After Sales Services. Springer, Berlin, Heidelberg, pp. 251–270.

[16] Becker, J., 2016. Dynamisches kennliniengestütztes Bestandsmanagement. Dissertation, Hannover.

[17] Schmidt, M., 2019. Modellgestützte Bewertung der Kapazitätsabstimmung im Umfeld veränderlicher Nachfrage. Dissertation, Hannover.

[18] Sendler, M., 2020. Ganzheitliche Konfiguration der PPS in der Instandhaltung hochwertiger Investitionsgüter. Dissertation, Braunschweig.

[19] Biedermann, H., 2008. Ersatzteilmanagement: Effiziente Ersatzteillogistik für Industrieunternehmen, 2., erw. u. aktualisierte Aufl. 2008 ed. Springer, Berlin, Heidelberg, 156 pp.

[20] Croston, J.D., 1972. Forecasting and Stock Control for Intermittent Demands. Journal of the Operational Research Society 23 (3), 289–303.

[21] Xu, Q., Wang, N., Shi, H., 2012. Review of Croston's method for intermittent demand forecasting, in: 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery. 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Chongqing, Sichuan, China. 29.05.2012 - 31.05.2012. IEEE, pp. 1456–1460.

[22] Syntetos, A.A., Boylan, J.E., 2005. The accuracy of intermittent demand estimates. International Journal of Forecasting 21 (2), 303–314.

[23] Teunter, R.H., Syntetos, A.A., Zied Babai, M., 2011. Intermittent demand: Linking forecasting to inventory obsolescence. European Journal of Operational Research 214 (3), 606–615.

[24] Romeijnders, W., Teunter, R., van Jaarsveld, W., 2012. A two-step method for forecasting spare parts demand using information on component repairs. European Journal of Operational Research 220 (2), 386–393.

[25] Ghobbar, A.A., Friend, C.H., 2003. Evaluation of forecasting methods for intermittent parts demand in the field of aviation: a predictive model. Computers & Operations Research 30 (14), 2097–2114.

[26] Hemeimat, R., Al-Qatawneh, L., Arafeh, M., Masoud, S., 2016. Forecasting Spare Parts Demand Using Statistical Analysis. AJOR 06 (02), 113–120.

[27] Voigt, J.P., 1973. Erfassung, Auswertung und Nutzung von Schadendaten in der Eisen- und Stahlindustrie. Dissertation, Braunschweig.

[28] Wilke, F.L., 1981. Bestimmung von Ausfällen und Störungen an Baugruppen des im Metallerzbergbau eingesetzten Fahrlader, Untersuchung ihrer Charakteristiken und Auswirkungen auf die Kosten des Gesamtbetriebes. Forschungsbericht Nr. 4397.

[29] Jacob, H., Hansmann, K.-W., Layer, M., Preßmar, D.B., Alex, B., 1998. Künstliche neuronale Netze in Management-Informationssystemen. Gabler Verlag, Wiesbaden.

[30] Gutierrez, R.S., Solis, A.O., Mukhopadhyay, S., 2008. Lumpy demand forecasting using neural networks. International Journal of Production Economics 111 (2), 409–420.

[31] Kozik, P., Sęp, J., 2012. Aircraft Engine Overhaul Demand Forecasting using ANN. Management and Production Engineering Review, 21–26.

[32] Şahin, M., Kızılaslan, R., Demirel, Ö.F., 2015. Forecasting Aviation Spare Parts Demand Using Croston based Methods and Artificial Neural Networks. Computer Science.

[33] Jensen, F.V., Nielsen, T.D., 2007. Bayesian networks and decision graphs, Softcover reprint of the hardcover 2nd ed. 2007 ed. Springer, Berlin, 447 pp.

[34] Boutselis, P., McNaught, K., 2019. Using Bayesian Networks to forecast spares demand from equipment failures in a changing service logistics context. International Journal of Production Economics 209, 325–333.

[35] Fu, W., Chien, C.-F., Lin, Z.-H., 2018. A Hybrid Forecasting Framework with Neural Network and Time-Series Method for Intermittent Demand in Semiconductor Supply Chain, in: Moon, I., Lee, G.M., Park, J., Kiritsis, D., Cieminski, G. von (Eds.), Advances in Production Management Systems. Smart Manufacturing for Industry 4.0, vol. 536. Springer International Publishing, Cham, pp. 65–72.

[36] Rosienkiewicz, M., Chlebus, E., Detyna, J., 2017. A hybrid spares demand forecasting method dedicated to mining industry. Applied Mathematical Modelling 49, 87–107.

[37] Pavya, K., Srinivasan, B., 2017. Feature Selection Techniques in Data Mining: A Study. International Journal of Scientific Development and Research 2 (6), 594–598.

[38] Arauzo-Azofra, A., Aznarte, J.L., Benítez, J.M., 2011. Empirical study of feature selection methods based on individual feature evaluation for classification problems. Expert Systems with Applications 38 (7), 8170–8177.

[39] Runkler, T.A., 2012. Data Analytics. Vieweg+Teubner Verlag, Wiesbaden.

[40] Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. The Journal of Machine Learning Research (3), 1532–4435.

[41] Hiller, T., Lucht, T., Kämpfer, T., Vinke, L., Holtsch, P., Nyhuis, P., 2021. Hybride Lieferzeitprognose: Verbesserte Termin- und Auftragsplanung im volatilen MRO-Umfeld. ZWF 116 (12), 882–888.

Biography

Torben Lucht, M.Sc. (*1991) studied industrial engineering with the focus on production technology at RWTH Aachen University. Since 2018, he works as a research associate in the field of production management at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover.

Volodymyr Alieksieiev, B.Sc. (*1999) studied engineering of logistic systems at National Technical University "Kharkiv Polytechnic Institute" (Ukraine). Since 2020 he studies production and logistics at Leibniz University Hannover. He also works as student assistant at the IFA since 2021.

Tim Kämpfer, M.Sc. (*1992) studied production engineering and logistics at Leibniz University Hannover and works as a research associate in the field of production management at the IFA at the Leibniz University Hannover since 2019.

Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research assistant at IFA. After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003. In 2008 he became managing partner of the IPH - Institut für Integrierte Produktion Hannover gGmbH.

3rd Conference on Production Systems and Logistics

Openness Of Digital Twins In Logistics – A Review

Stephanie Winkelmann[1], Hendrik van der Valk[1]
[1]*Chair for Industrial Information Management, TU Dortmund University, Dortmund, Germany*

Abstract

Openness is becoming increasingly important in scientific research and practice. It describes the phenomenon of sharing information with other internal or external stakeholders by using different technologies, e.g., cloud computing, distributed ledger, or digital twins. Hence, many researchers investigate and evaluate the openness of platforms. Alongside these platforms, digital twins are gaining influence in industrial processes. A digital twin is a virtual representation of a physical entity connected through a bi-directional data linkage. Its primary purpose is to visualize, analyze, and optimize production and logistics systems. Nevertheless, research shows a lack of knowledge in the domain of the openness of digital twins and that the topic has not been addressed adequately. To approach this research gap, this paper provides a review of literature-based work on digital twins focusing on logistical contexts. It aims to answer the question of how open digital twins are, depending on their use case, purpose, and status as digital twin or digital shadow. Through a comprehensive research approach, this paper provides researchers and practitioners with meaningful insights into the openness of digital twins.

Keywords

Openness; Digital Twin; Production and Logistics; Review

1. Introduction

It was inconceivable to make internal operations transparent to external stakeholders for decades. Even within a supply chain, each company operated within its premises. But transparency along a supply chain holds many advantages for logistics because logistics connects places and companies in global networks and creates value [1]. Therefore, the aspect of openness has become increasingly important, especially in research. Openness results from transparency, which in turn is created by the exchange of data between different entities [2]. This data exchange is supported, among other things, by a so-called digital twin. The digital twin is a virtual construct of an actual entity with a bidirectional connection [3]. It is this connection that enables new applications. Following [4], the interest in the digital twin has increased in research and industry. Many companies see great potential in using the digital twin [2], so it will gain further influence on industrial operations in the future. One of the primary purposes of a digital twin is to create transparency in logistics by solving problems regarding visibility [5]. Nevertheless, as with many digital constructs, e.g., virtual platforms, data sharing is often limited by user requirements, data sovereignty of owners, as well as suppliers [6]. At this point, the following research question arises:

RQ1: Are digital twins enabling data sharing within ecosystems?

DOI: https://doi.org/10.15488/12186

Before research may address the level of openness, we must ensure that digital twins, in principle, can provide transparency through data sharing. Then, if data sharing capability is guaranteed, we examine the level of openness of digital twins for data sharing. Thus, the second research question reads as follows:

RQ2: How open are digital twins used in logistics?

To answer the research questions, we conduct an exploratory examination of the topic area based on a systematic literature review following [7] as well as [8]. The paper is structured as follows. First, we provide insight into the concepts of digital twins, their data sharing capabilities, and the general concept of openness. Then, we give an overview of the research method before we describe the literature review results. After that, we explain and discuss our observations. Finally, we summarize the findings and offer contributions, limitations, and an outlook for further research topics.

2. State of the Art

2.1 Digital Twins

Digital twins originate from the decades-old concept of physical twins as simulation and experiment environments for real-world applications [4]. One of the first noted deployments was the Apollo Space Project, in which physical twins mirrored the space capsules for testing purposes. Since then, the concept of digital twins has developed continuously. Starting from [3], who defined the digital twin as a combination of physical and virtual products that are connected by data and information. Later, [9] extended their view as they now see sensors as the primary data source for digital twins. They laid the ground for the digital twin as a concept for product life cycle management [10].

Simultaneously, NASA pushed digital twin concepts further. Researchers from this ecosystem see the digital twin as "an integrated multiphysics, multiscale, probabilistic simulation" [11, p.7]. This paves the way for the second research stream in which the digital twin is seen as the current development stage of the classical simulation of production and logistic processes [4]. Both research streams combine the fact that they lack a fundamental understanding of what a digital twin consists of [12]. In the last two years, extensive research and work have tackled this research gap. [13] concentrate on the data flows towards and from a digital twin to specify the concept. They demand a bi-directional data flow for digital twins as the distinguishing feature between digital twins and digital shadows as well as digital models. [14] extends the digital twin by the dimension of services and examines the question of what a digital twin should be able of in hindsight to data processing. A more thorough analysis was the creation of a taxonomy of digital twins, which describes the digital twin in the eight dimensions data link, purpose, conceptual elements, accuracy, interfaces, synchronization, data input, and time of creation [15]. [16] come to similar but more nuanced dimensions with their twelve characteristics and extend the eight dimensions with the aspects of physical and environment, fidelity, and system state. The latest development is five archetypes of digital twins, ranging from a basic twin with low capabilities to a fully enhanced digital twin that can process complex operations and monitor and control physical systems [17]. From this research, we follow the most recent and concluding definition: "The digital twin is a virtual construct that represents a physical counterpart, integrates several data inputs with the aim of data handling, data storing, and data processing, and provides an automatic, bi-directional data linkage between the virtual world and the physical one. Synchronization is crucial to the digital twin to display any changes in the state of the physical object. Additionally, a digital twin must comply with data governance rules and must provide interoperability with other systems" [17, p. 14]. Especially, the new and not yet in combination with digital twins portrayed dimensions of data governance and interoperability are crucial for this paper. Data governance may consist of many rules, but it is not specified which rules should apply. For interoperability, three configurations are provided by [17]. There is either no interoperability, a certain degree of interoperability via translation devices within the interfaces, or full interoperability between all agents within a given ecosystem. We assume a high level of openness for a fully interoperable digital twin, which we will analyze in this paper. A related concept to digital twins is the concept of digital shadows. In this paper, we follow the differentiation of [13]. Hence, a digital shadow shows many aspects and properties of a digital twin but lacks a bi-directional, automatic data flow.

[18] have investigated the use of the digital twin and provide an overview of the industries in which digital twins are used. Here, the Digital Twin supports simulation, monitoring as well as optimization of the physical plant [18,15].

2.2 Openness

The term *openness* describes the way technologies are used concerning exchanges with other stakeholders. The focus here is on the use of technology [19]. According to [20], openness is achieved through collaboration between different actors. It is possible that organizational boundaries may limit the actors. Thus, it is unnecessary to share the technology with external stakeholders to operate "open". Thereby, transparency is one prerequisite for openness [21]. The challenge is to find the right balance between control and openness [22–25]. One challenge, for example, is to create governance rules that appropriately limit participants' freedom of action [22]. Therefore, [19] sort openness into different degrees. Closed technologies have the smallest degree of openness. Only one actor controls them. Usually, the access for other actors is restricted by the owner through, e.g., the imposition of patents or copyrights [19,26]. Opposing technologies are those that are used to be purely open. They are accessible to all actors [19]. Between both extremes are many levels of openness, which depend on individual use cases.

The combination of openness and information technologies generates opportunities. Information technologies enable a broad scope of open practices such as open source, open source software, and open innovation [27]. Open source is often used in the field of programming. The idea is to unify the efforts of programmers. Sharing the code or granting access to the code are essential parts of being open source in programming. The term is based on open source software [28,29]. The unique feature of the open source software is that a comprehensive group of users has access to the software's source code. Furthermore, they are allowed to use and to modify the software. These changes give rise to further artifacts, which in turn can be distributed.

In recent years, the term *open innovation* achieved a lot of attention. Open innovation describes the necessity to access a technology and concentrate on an open research and development process [30]. [31] has developed six principles regarding open innovation. Among other things, he assumes that it is crucial to use the expertise and experience of external parties. This extern research and development create an added value that could be useable for the intern analysis. The optimum will be achieved by combining the intern and the extern research and development. Since there are several definitions of the term open innovation, [32] defines it newly. According to him, open innovation is a distributed innovation process. The basis of this process is a consortium of precisely controlled knowledge flows. These go beyond the boundaries of the company. These knowledge flows are used in a targeted manner in line with the business model.

3. Research Method

The aim is to analyze the digital twin in logistics in terms of its openness. For this reason, the research method starts with a structured literature analysis according to [7] and [8] to obtain an overview of existing literature. Thus, the first step is the database search. Therefore, the search string has been defined. The search should focus on the digital twin, which results in the first part of the string *"Digital Twin"*. Furthermore, the search is to be restricted to the domain of logistics. Since in English, the terms "logistics" and "supply chain" are sometimes used as synonyms, the second part of the search string *("Logistics" OR "Supply Chain")* results. The subject area of openness is not specifically narrowed down at this point. Some publications describe the construct of openness but do not explicitly call it openness. The search result confirms this statement (see [33,34]). The entire search string reads as follows *"Digital Twin" AND ("Logistics" OR "Supply Chain")*. This has been entered into five common databases (AISeL, IEEEXplore, WoS, Scopus, and Science Direct). To make possible developments of the digital twin visible, the search has not been restricted to a publication period. Also, to obtain a high-quality literature review, the search has been limited to peer-reviewed papers. Additionally, only papers in English have been included. The period in which the search has been carried out extends through winter 2021/2022.

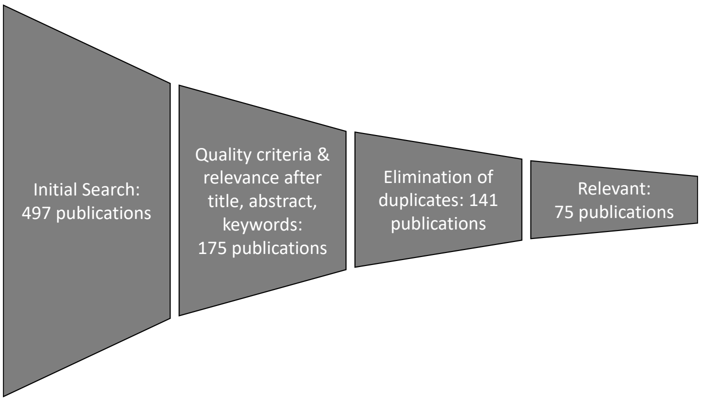


Figure 1: Search Process.

The search, including forward and backward searches, resulted in 497 publications (see Figure 1), of which 175 were declared relevant after an analysis of title, abstract, and keywords and the application of the quality criteria. As quality criteria, we demanded a peer-review process, a sound description of the research process, the application of commonly accepted research methods, and a noticeable consistency throughout the paper. The extraction of duplicates resulted in a literature sample of 141 publications. These were entirely analyzed by the two authors and then included in the final literature sample according to their relevance. The final sample consists of 75 publications.

4. Openness of Digital Twins

To answer how open digital twins are, we first have to ensure the overall capability of a digital twin to be open (see RQ1). We define a concept as being able to be open if it can provide transparency over a process. [17] state that the two most important purposes of a digital twin are simulation and monitoring. Deriving from there, we may expect the capability to provide transparency. In fact, monitoring any process will bring transparency to the monitored process. Additionally, simulation operations offer an overview of a particular system and provide transparency. [35] even developed an architectural model for digital twins primarily used to create transparency in supply chains based on the International Data Spaces and their connectors. As operational digital twins in production and construction already exist (e.g., [36], [37]), which create transparency, we attest to the digital twin`s ability to create transparency.

This leads to the second research question, how open are digital twins in logistics. During the analysis, we could identify three classifications of openness: intraorganizational, dual, and multisided. If the digital twin is just implemented within one participant of the supply chain, we allocate the digital twin to the label intraorganizational. There is no exchange of information between different participants of the supply chain via the digital twin. If the digital twin is shared with only one other participant in the supply chain, we label the digital twin as dual. The third label is called multisided. This label describes the implementation of the digital twin by more than one participant of the logistical network. Data and information are shared within this ecosystem. These three classes align with the postulation that there are different degrees of openness (see section 2).

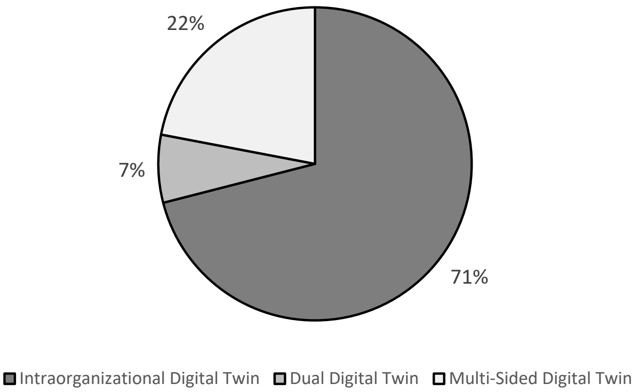


Figure 2: Shares of Openness.

An analysis of our concept matrix shows that 71% of the articles describe the implementation of the digital twin within just one participant (see Figure 2). The other 29% percent are split between dual (7%) and multisided (22%). This result contradicts our expectations that the participating enterprises share the digital twin within the entire supply chain. An explanation for this is that enterprises may see risks regarding data security, governance, or data abuse [38]. Therefore, they use the digital twin primarily within their own enterprise.

To gain a deeper insight, we specify the analysis and search for relationships between the openness and other categories like use case, purpose, and twin type. The use case is based on the established Supply Chain Operations Reference Model (SCOR Model). This model is used for standardizing the processes in a supply chain. It consists of the five different operational phases: plan, source, make, deliver, and return [39]. Additionally, we note that digital twins have three primary logistics purposes: simulation, optimization, and monitoring [17]. Furthermore, there are two twin types, digital twin and digital shadow, which are explained in section 2.1. Table 1 visualizes the results.

Table 1: Correlations Openness (max. values highlighted).

Meta-Dimension	Dimension*	Intraorganizational	Dual	Multisided
Use Case	Plan	18%	17%	11%
	Source	7%	**33%**	11%
	Make	**61%**	25%	22%
	Deliver	12%	25%	**49%**
	Return	2%	0%	7%
Purpose	Simulation	**46%**	28%	**52%**
	Optimization	26%	**43%**	16%
	Monitoring	28%	29%	32%
Twin Type	Digital Twin	**83%**	**67%**	**85%**
	Digital Shadow	17%	33%	15%

*Each object may address multiple dimensions within one meta-dimension. However, only the primary dimension is noted.

Due to space limitations, we have not included the conceptual matrix with all 75 publications in this paper. The results show some interesting insights. Starting with the use cases, we notice different distributions between the level of openness. Intraorganizational digital twins show the most significant variations between the extreme values. Sixty-one percent of the analyzed digital twins are mainly operated in the make phase. Hence, this is the most common relation. As the make phase is often within one enterprise, the usage of intraorganizational twin is plausible. In these cases, the digital twin is mainly used to simulate and optimize the enterprise's own material flow and not the entire supply chain. On the contrary, just two percent of the digital twins are used within the return phase. These cases often include the former customer and an external return company. The manufacturer is mostly not integrated into these processes. Hence, intraorganizational twins are demanding, as this is a highly transactional business that includes at least three parties (manufacturer, customer, and return company).

Dual twins are most common in the phase source, in which bilateral relationships between manufacturer and supplier are common. Though, make and deliver are essential phases for dual twins as well. Whereas multisided digital twins are concentrated on the deliver phase. This is no surprise, as a manufacturer naturally delivers its products to many customers who need access to the digital twin. That nearly one-quarter of the dual and multisided digital twins are used in the make phase shows the potential of the more open twins. The production processes are more and more intertwined with other factories and companies. Hence, a digital representation, which follows this exchange level, needs to be in place.

The main benefit of a digital twin is the possibility of simulation [40]. So, it poses no surprise that many digital twins for simulation purposes are intraorganizational twins. However, it is surprising that the dual twins` primary purpose is optimization. Often these twins optimize logistics flows between two companies. So, these two companies are forced to share data. Nevertheless, often they do not want to integrate further participants. Hence, multisided twins are not so common in optimization.

These multisided digital twins are more often used for visibility, monitoring, and simulation purposes. Visibility is justified by the aim of transparency across the entire supply chain. Hence, data from various suppliers need to be integrated. Similarly, simulation gets better as the amount of data increases. So, it is reasonable that the consolidation of data from different enterprises supports the simulation. As expected, most analyzed objects describe true digital twins. The intraorganizational and multisided digital twins are mostly true twins, as the commonly related twin type is the digital twin. However, many so-called dual digital twins are digital shadows (33%), following the definition of [13]. Hence, the digital shadow is common when shared with other enterprises. As mentioned, enterprises see risks in sharing their data with other enterprises. Through the manual data return flow provided by a digital shadow, the enterprise can influence the quantity of the shared data. Using a digital twin will lose this influence because the digital twin shares the data automatically with the other enterprises. In addition, the automatic data return flow could directly influence the enterprise's productivity when the other enterprise performs changes via the digital twin.

5. Conclusion and Outlook

This paper reviews the levels of openness of digital twins in the domain of logistics. Therefore, we analyzed the literature through a structured literature review. Regarding RQ1, we state that digital twins are per se able for interorganizational data sharing. They are able to create transparency over a process. Hence, digital twins provide the basics and fundamental capabilities to be considered open. For RQ2, the review gives us a deeper look at the openness of digital twins in logistics. The results show that combining the topics of digital twin and openness opens a new field of research.

Regarding openness, the focus on intraorganizational digital twins is rather astonishing. This result contradicts our expectations that the participating enterprises share the digital twin within an entire supply chain. An explanation for this is that enterprises may see risks regarding data security, data governance, or data abuse. Through the manual data return flow, which is provided by a digital shadow, the participants have the possibility to influence the quantity of the shared data. If they use a digital twin, they may lose this influence because the digital twin should share the data automatically with other participants. In addition, the automatic data return flow could directly influence the participant's productivity when another participant performs changes via the digital twin. Therefore, they use the digital twin primarily within their own enterprise. At this point, a research gap arises since openness is an important topic in logistics. Hence, a certain level of visibility is justified by the aim of transparency across the entire supply chain. Concluding the discussion, we identify several research gaps that should be investigated in further research:

- Focus on the openness of digital twins in practice
- Providing reference architecture and standard procedures for data security in open digital twins
- Description of industrial applications and use cases
- Focus on additional domains besides classical logistics

Our work is subject to certain limitations. Even if we tried to keep any subjective influence to a minimum, the classification naturally suffers subjective influences. In particular, other researchers may make the distinction between digital twins and the evaluation of openness differently. Therefore, they obtain different results. However, this research makes multiple contributions to the corps of scientific research, as well as managerial contributions. We structure the literature on digital twins in logistics and visualize certain conclusions regarding the openness of digital twins. These conclusions provide white spots that are suitable for further research. Hence, new research streams may be implemented upon this paper. The managerial contributions are not quite as direct as the scientific ones. Digital twins still lack a broad operation base in logistics contexts. Nevertheless, we provide the industrial experts with the prerequisite for a deeper look at their projects regarding open digital objects that accompany their logistics. Furthermore, practitioners will benefit from future research on this topic.

Acknowledgements

This research was conducted in the project "Silicon Economy Logistics Ecosystem" and is funded by the German Federal Ministry of Transport and Digital Infrastructure.

References

[1] Barreto, L., Amaral, A. & Pereira, T. (2017). Industry 4.0 implications in logistics: an overview. Procedia Manufacturing, 13 (1), 1245–1252.

[2] Whittington, R., Cailluet, L., Yakis-Douglas, B., 2011. Opening Strategy: Evolution of a Precarious Profession. British Journal of Management 22 (3), 531–544.

[3] Grieves, M., 2014. Digital Twin: Manufacturing Excellence Through Virtual Factory Replication. Michael W. Grieves LLC.

[4] Rosen, R., Wichert, G. von, Lo, G., Bettenhausen, K.D., 2015. About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. IFAC-PapersOnLine 48 (3), 567–572.

[5] Haße, H., van der Valk, H., Weißenberg, N., Otto, B., 2020. Shared Digital Twins: Data Sovereignty in Logistics Networks, in: Proceedings of the HICL.

[6] Otto, B., Jarke, M., 2019. Designing a Multisided Data Platform: Findings From the International Data Spaces Case. Electronic Markets 43 (1), 39.

[7] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), xiii–xxiii.

[8] Vom Brocke, J., Simons, A., Niehaves, B., Reimer, K., Plattfaut, R., Cleven, A., 2009. Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process, in: Proceedings of the 17th European Conference on Information Systems. AIS, Verona.

[9] Grieves, M., Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems, in: Kahlen, F.-J., Flumerfelt, S., Alves, A. (Eds.), Transdisciplinary Perspectives on Complex Systems. New Findings and Approaches. Springer International Publishing, Cham, Switzerland, pp. 85–113.

[10] Wagner, C., Grothoff, J., Epple, U., Drath, R., Malakuti, S., Gruner, S., Hoffmeister, M., Zimermann, P., 2017. The Role of the Industry 4.0 Asset Administration Shell and the Digital Twin During the Life Cycle of a Plant, in: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation. IEEE, Piscataway, USA, pp. 1–8.

[11] Glaessgen, E., Stargel, D., 2012. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles, in: Structures, Structural Dynamics, and Materials and Co-located Conferences. 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference. American Institute of Aeronautics and Astronautics, Reston, USA, 1-14.

[12] Haag, S., Anderl, R., 2019. Automated Generation of As-Manufactured Geometric Representations for Digital Twins Using STEP. Procedia CIRP 84, 1082–1087.

[13] Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51 (11), 1016–1022.

[14] Tao, F., Zhang, H., Liu, A., Nee, A.Y.C., 2019. Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics 15 (4), 2405–2415.

[15] van der Valk, H., Haße, H., Möller, F., Arbter, M., Otto, B., 2020. A Taxonomy of Digital Twins, in: , AMCIS 2020 Proceedings. AIS, Salt Lake City, USA.

[16] Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., 2020. Characterising the Digital Twin: A Systematic Literature Review. CIRP journal of manufacturing science and technology 29, 36–52.

[17] van der Valk, H., Haße, H., Möller, F., Otto, B., 2021. Archetypes of Digital Twins. Business & Information Systems Engineering.

[18] Enders, M.R., Hoßbach, N., 2019. Dimensions of Digital Twin Applications - A Literature Review, in: Proceedings of the 25th Americas Conference on Information Systems, Cancun: Mexico, pp. 1–10.

[19] Boudreau, K., 2010. Open Platform Strategies and Innovation: Granting Access vs. Devolving Control. Management Science 56 (10), 1849–1872.

[20] Dokukina, A.A., Petrovskaya, I.A., 2020. Open Innovation as a Business Performance Accelerator: Challenges and Opportunities for the Firms' Competitive Strategy, in: Solovev, D.B., Savaley, V.V., Bekker, A.T., Petukhov, V.I. (Eds.), Proceeding of the International Science and Technology Conference "FarEastCon 2019", vol. 172. Springer Singapore, Singapore, pp. 275–286.

[21] Ågerfalk, P.J., Fitzgerald, B., Stol, K.-J., 2015. Innersourcing, in: Ågerfalk, P.J., Fitzgerald, B., Stol, K.-J. (Eds.), Software Sourcing in the Age of Open, vol. 52. Springer International Publishing, Cham, pp. 27–44.

[22] Constantinides, P., Henfridsson, O., Parker, G.G., 2018. Platforms and Infrastructures in the Digital Age. Information Systems Research 29 (2), 381–400.

[23] Fuerstenau, D., Rothe, H., Baiyere, A., Schulte-Althoff, M., Masak, D., Schewina, K., Anisimova, D., 2019. Growth, Complexity, and Generativity of Digital Platforms: The Case of Otto.de, in: Proceedings of the 40th International Conference on Information Systems, Munich: Germany.

[24] Hein, A., Schreieck, M., Riasanow, T., Setzke, D.S., Wiesche, M., Böhm, M., Krcmar, H., 2020. Digital platform ecosystems. Electronic Markets 30 (1), 87–98.

[25] Parker, G., van Alstyne, M.W., Jiang, X., 2016. Platform Ecosystems: How Developers Invert the Firm. SSRN Electronic Journal 31 (6), 305.

[26] Cohen, W., Nelson, R., Walsh, J., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). National Bureau of Economic Research, Cambridge, MA.

[27] Schlagwein, D., Conboy, K., Feller, J., Leimeister, J.M., Morgan, L., 2017. "Openness" with and without Information Technology: A Framework and a Brief History. Journal of Information Technology 32 (4), 297–305.

[28] Fitzgerald, 2006. The Transformation of Open Source Software. MIS Quarterly 30 (3), 587.

[29] Morgan, L., Finnegan, P., 2014. Beyond free software: An exploration of the business value of strategic open source. The Journal of Strategic Information Systems 23 (3), 226–238.

[30] Trott, P., Hartmann, D., 2009. Why 'Open Innovation' is old wine in new bottles. International Journal of Innovation Management 13 (04), 715–736.

[31] Chesbrough, H.W., 2003. Open innovation: The new imperative for creating and profiting from technology, [Nachdr.] ed. Harvard Business School Press, Boston, Mass., 227 pp.

[32] Chesbrough, H., 2017. The Future of Open Innovation. Research-Technology Management 60 (6), 29–35.

[33] Barykin, S.Y., Bochkarev, A.A., Kalinina, O.V., Yadykin, V.K., 2020. Concept for a Supply Chain Digital Twin. International Journal of Mathematical, Engineering and Management Sciences 5 (6), 1498–1515.

[34] Defraeye, T., Shrivastava, C., Berry, T., Verboven, P., Onwude, D., Schudel, S., Bühlmann, A., Cronje, P., Rossi, R.M., 2021. Digital twins are coming: Will we need them in supply chains of fresh horticultural produce? Trends in Food Science & Technology 109 (1), 245–258.

[35] Cirullies, J., Schwede, C., 2021. On-demand Shared Digital Twins – An Information Architectural Model to Create Transparency in Collaborative Supply Networks, in: Proceedings of the 54th Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences.

[36] Zhao, R., Yan, D., Liu, Q., Leng, J., Wan, J., Chen, X., Zhang, X., 2019. Digital Twin-Driven Cyber-Physical System for Autonomously Controlling of Micro Punching System. IEEE Access 7, 9459–9469.

[37] Opoku, D.-G.J., Perera, S., Osei-Kyei, R., Rashidi, M., 2021. Digital twin application in the construction industry: A literature review. Journal of Building Engineering 40, 102726.

[38] Jarke, M., Otto, B., Ram, S., 2019. Data Sovereignty and Data Space Ecosystems. Business & Information Systems Engineering 61 (5), 549–550.

[39] Association for Supply Chain Management, 2017. Supply Chain Operations Reference Model (SCOR): Version 12.0.

[40] van der Valk, H., Hunker, J., Rabe, M., Otto, B., 2020. Digital Twins in Simulative Applications: A Taxonomy, in: Bae, K.-H., Feng, B., Kim, S., Lazarova-Molnar, S., Zheng, Z., Roeder, T., Thiesing, R. (Eds.), Proceedings of the 2020 Winter Simulation Conference. IEEE, Piscataway, NJ, USA, pp. 2695–2706.

Biography

Stephanie Winkelmann (*1994) has been a researcher at the Chair for Industrial Information Management at the TU Dortmund University since 2022. She has graduated with a Bachelor's degree in Business Administration and Logistics and holds a Master of Science in Logistics from TU Dortmund University. In cooperation with the Fraunhofer Institute for Software and Systems Engineering and the Fraunhofer Institut for Material Flows and Logistics, she conducts her research as a member of the Silicon Economy.

Hendrik van der Valk (*1992) has been a researcher at the Chair for Industrial Information Management at the TU Dortmund University since 2019. He studied Mechanical Engineering, and his research focuses on Digital Twins in production and logistics. He conducts his research as a member of the Excellence Center for Logistics and IT and the Silicon Economy. He is currently a member of the Association for Information Systems AIS and the Institute for Operations Research and the Management Sciences INFORMS.

3rd Conference on Production Systems and Logistics

Adaption Of The Level Of Development To The Factory Layout Planning And Introduction Of A Quality Assurance Process

Thomas Neuhäuser[1,2], Lisa Lenz[3], Kai Weist[3], Andrea Hohmann[1], Rüdiger Daub[1,2]
[1]*Fraunhofer Institute for Casting, Composite and Processing Technology*
[2]*Institute for Machine Tools and Industrial Management, Technical University of Munich*
[3]*Department of Construction Management, TU Dortmund University*

Abstract

Current developments and trends are causing an increasingly turbulent environment for manufacturing companies. In order to respond to these dynamic market conditions, products and thus also production systems have to be adapted more frequently and much faster. However, time and cost targets are often missed by classic factory planning approaches due to poor communication, inadequate tools, and lack of interfaces. Therefore, new ways have to be found in factory planning to overcome these problems. Building Information Modeling, which is already used in the construction industry, provides a promising method for the collaboration of stakeholders based on digital models. This would allow communication to be structured, new tools to be used, and interfaces to be stabilized to improve the target achievement in factory planning projects. However, which information should be provided in which level of detail in which phase of a factory planning project and how the quality of this information can be ensured has not yet been answered. A possible solution to these questions is addressed in this article. First, the concept of the so-called Level of Development, i.e. the geometric and non-geometric definition of the model contents, is transferred to factory layout planning. Then, based on two use cases, the process of quality assurance is defined.

Keywords

Factory Planning; Building Information Modeling; Level of Development; Layout Planning

1. Introduction

The manufacturing industry is one of Germany's most important economic sectors, accounting for 23.5% of gross domestic product [1]. However, companies in this sector are facing trends such as globalization, dynamization of product life cycles and climate change as well as current challenges such as supply bottlenecks. These new circumstances force manufacturing companies to adapt and innovate. The focus is on new products, new processes, innovative supply networks and also the adaptation of existing or the creation of new factories. The adaptation cycles must be implemented faster and more frequently. Factory planning is thus becoming a continuous task for companies. [2–4]

The central target variables of planning projects are time, costs and quality. While the quality targets are generally achieved in factory planning projects, the time targets are missed in about 60% and the cost targets in approximately 72%. This is primarily due to four areas of potential improvement in the organization of factory planning projects: An improvement of communication and synchronization, a further development of instruments and tools used, an increase in planned agility in the course of the project, and an early

DOI: https://doi.org/10.15488/12137

detection of deviations [5]. The costs of errors on German construction sites amounted to 18.3 billion Euros in 2020 [6]. To solve these problems and leverage the potential for improvement, new approaches to factory planning must be found.

It is precisely this potential that the Building Information Modeling (BIM) methodology addresses. BIM is a collaborative working methodology based on digital models, called the Building Information Model. The Building Information Model is the primary instrument of the methodology, which is generated by tools such as authoring software. Those involved in the planning process regularly exchange the models in a systematic communication process, which are checked for deviations at an early stage and thus continuously synchronized. Agility corridors are planned into the project from the beginning in order to be able to eliminate any deviations and, if necessary, to draw several iteration loops. In addition, each project is built up modularly via project-specific goals and use cases [7–9]. Current studies show that the use of BIM in construction projects leads to a reduction in time in 34% and to a reduction in costs in 60% of the investigated cases [10].

While there are recommendations for the use of the BIM methodology in public and municipal construction as well as in infrastructure construction [8,9,11], the use of the methodology in factory planning is still largely unexplored [12,13]. To overcome these shortcomings, this paper presents a modeling guideline regarding the geometric and non-geometric level of detail in the factory planning process. Then, two BIM use cases to ensure modeling quality are presented and finally described as a process. Thereby, the focus is laid on the layout planning process, since in this phase the production system merges with the building to form the factory and a large part of the planning interaction between the individual trades takes place. The approach presented will be primarily geared towards German companies, as the *Honorarordnung für Architekten und Ingenieure (HOAI)* as well as the *VDI-Richtlinie 5200* will be used as basis [14,15]. For this purpose, Chapter 2 explains the fundamental procedure for the application of the BIM method. Chapter 3 shows the different modeling contents of the planning phases, the so-called Level of Development (LoD). In Chapter 4 the process of quality assurance is discussed. Finally, the paper ends with a summary and a conclusion.

2. Fundamentals

2.1 Application of the BIM method

In order to apply the BIM method in organizations and projects, there are already initial recommendations for action [8,9,11]. The procedure is divided into three steps, the formulation of BIM goals, BIM use cases and BIM processes. The BIM goals, i.e., which results are expected through the application of the BIM method, serve as the starting point. Examples of this are the increase in planning quality, cost and time certainty. The BIM use cases are then defined. These concretize the BIM goals as activities. To increase the planning quality, for example, the two use cases geometric collision checking between partial and functional models and quality checking of the models (cf. Chapter 4) can be formulated. The BIM processes, which thus describe the actual working method with BIM, can then be derived from these use cases. This procedure is illustrated in Figure 1. [9,16]

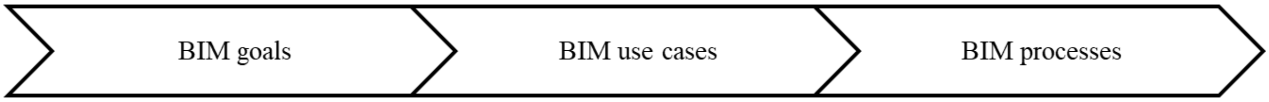


Figure 1: Procedure for the application of the BIM method according to [9,16]

When applying the BIM method in the context of projects, a distinction must also be made between little bim and BIG BIM, and between open and closed BIM (cf. Figure 2). Little bim is the use of BIM software products as a so-called isolated solution. For example, one planner uses BIM software, while the other participants use conventional software products. The opposite of little bim is BIG BIM. Here, all parties involved in the planning process work with digital building models according to the BIM method. The terms open and closed BIM refer to the exchange of data between the stakeholders. In a closed BIM approach, only the software of a specific manufacturer and its proprietary data exchange format is used. In contrast, in an open BIM approach, the software can be freely selected and the manufacturer-independent Industry Foundation Classes (IFC) format is used for data exchange and the BIM Collaboration Format (BCF) for model-based collision communication. [7]

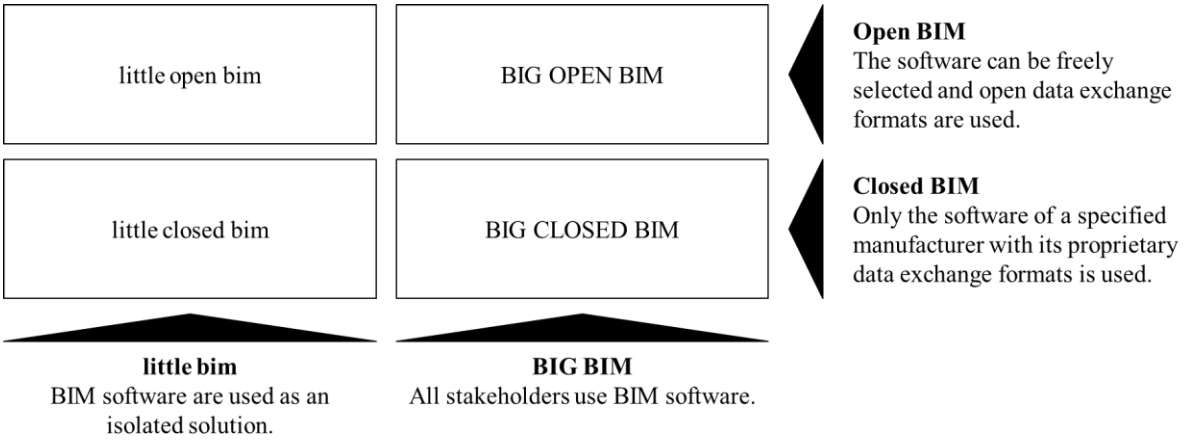


Figure 2: Breadth and data exchange of the application of BIM according to [7]

2.2 Development of the factory layout

Factory layout planning describes the spatial arrangement of operational structural units. The factory is understood as a system in which individual elements are related to each other. The relationships form the structure of the factory, while the elements represent the individual structural units. Depending on the level of factory planning, the structural units can have different characteristics, resulting in two different types of layout with increasing levels of detail, as shown in Figure 3. [4,17]

The first step of factory layout planning is the set-up of rough layouts (ideal and real), in which the functional areas (especially production and logistics areas) within the factory building are shown together with the main transport routes. Subsequently the fine layout is developed, in which the building services and media supply are planned in detail and the operating equipment is precisely positioned. Operating resources are defined as technical systems, equipment and facilities that are used to implement manufacturing, assembly and logistics processes. From a factory planning perspective, operating resources are production resources such as manufacturing machines, assembly resources such as joining tools, and logistics resources such as packaging equipment. All machines, tools, materials as well as energy and media connections are represented and thus the microstructure of the factory is visible. [4,15,17]

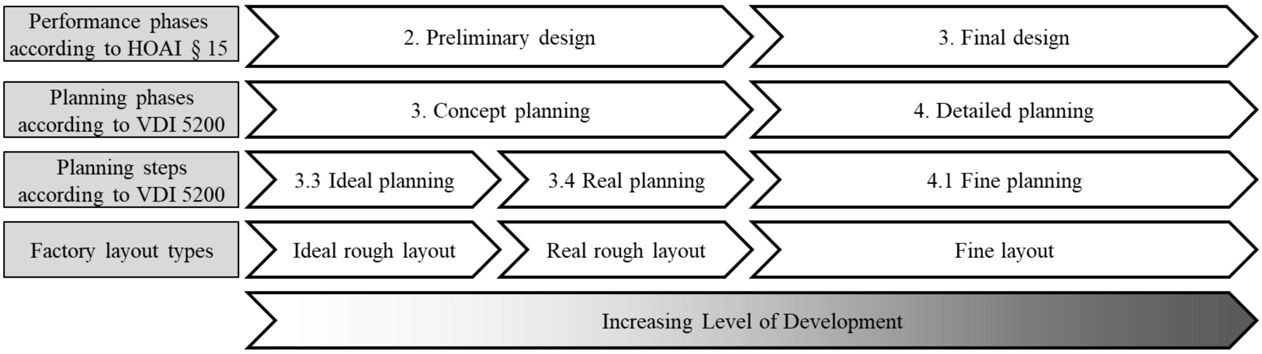


Figure 3: Development of the factory layout during the planning process according to [4,14,15,17]

In contrast to Grundig's explanations [17], the third layout type, the workstation layout, is explicitly dispensed with and its contents are integrated into the fine layout. The layout types are thus adopted according to *VDI-Richtlinie 5200* [15] since increasing the modeling detail within a planning phase is suboptimal for the quality assurance process. This would result in two different levels of modeling detail for this planning phase, and it would not be possible to clearly delineate according to which one quality assurance is to be performed. In addition, the rough layout is separated into ideal and real layout in order to be able to distinguish the content development of the modeling detail in the steps of ideal and real planning in Chapter 3.

Such a development of the modeling detail over time is described in the BIM method with the Level of Development (LoD). The LoD is composed of the geometric modeling level, the Level of Geometry (LoG) and the non-geometric attribution level, the Level of Information (LoI). The literature distinguishes between five general LoDs with possible intermediate levels, 100 to 500, which show the increasing complexity in terms of qualitative granularity (accuracy) and quantitative granularity (richness of detail) of geometric and non-geometric information over the course of a project. However, these levels largely refer to elements from architecture, technical building equipment and structural design and are not transferred to the development of the factory layout yet. An approach for this is presented in the next Chapter 3. [7,18]

3. Level of Development in the phases of factory layout planning

If the development of the factory layout representation is described using the LoD, this results in two levels for the concept as well as the detailed planning. In terms of the LoD in the building design process, these phases run parallel to performance phases 2 and 3 according to the HOAI (cf. Figure 3), resulting in LoD 100 to 200 for the layout planning phases [14,15,19]. The first two layout types, the ideal and the real rough layout, are both to be regarded as a preliminary draft model with LoD 100 and the final fine layout as the draft model with LoD 200. This is shown in Table 1.

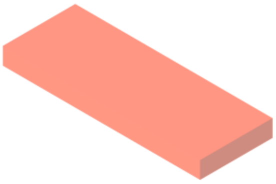
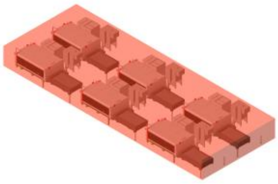
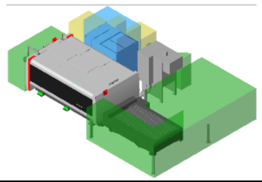
The ideal rough layout is represented as a three-dimensional block layout in which the functional areas of production and logistics are located. The blocks reflect the approximate space requirements and the main route network is also shown. Functional descriptions of the blocks are integrated as non-geometric information and special building design requirements are defined.

The real rough layout is still modeled as a three-dimensional block layout on the area or segment level [15], but the LoG can already be extended to include machine drawings or envelope models. However, care must be taken not to use detailed Computer Aided Design (CAD) design models of machinery and facilities, as these can significantly affect the performance of the overall model. As non-geometric information, static and dynamic loads, required media, caused emissions and lighting requirements should also already be assigned to the individual blocks. If necessary, the blocks can be further subdivided, leaving the area level.

In the final fine layout, the individual resources and workstations are shown in detail. The LoG is clearly deepened with the exact arrangement of the operating equipment, the representation of media connection points as well as maintenance and servicing, logistics and work spaces. In addition, the non-geometric information previously available at area level is assigned to the individual operating equipment and localized there.

Without the clear definition of the LoD for all specialist planners at the start of the project, there is no comparability of the models. The LoD should therefore be an integral part of the Employers Information Requirements (EIR), respectively, the BIM Execution Plan (BEP) in each factory planning project in order to be able to fulfill the BIM goals and use cases based on the Building Information Model. Thus, the LoD is an essential basis for the quality assurance process described in the following chapter.

Table 1: Level of Development in the phases of factory layout planning according to Hausknecht et al. [20]

	LoD 100	**LoD 100**	**LoD 200**
	Ideal rough layout	Real rough layout	Fine layout
Illustration			
Description	First ideal arrangement of production and logistics areas as preliminary draft model	Investigation of layout variants and determination of the preferred variant as preliminary draft model	Detailed layout up to the representation of the individual operating resources and workstations as draft model
LOG	Three-dimensional block layout with the main route network of the factory with approximate space requirements	Three-dimensional block layout with the main route network of the factory with machine drawings or envelope models	Detailed arrangement of equipment, representation of media connection points as well as, maintenance and service, logistics and work spaces with specific dimensions, location and orientation
LOI	Functional description of the blocks including special requirements	First general alphanumeric information such as static and dynamic loads, required media and causing emissions as well as lighting requirements	Detailing of all alphanumeric information, such as the assignment of media connection values to the media connection points

4. Assurance of the planning quality

When it comes to assurance of planning quality, the focus is primarily on the BIM goal of increasing planning quality. However, if this BIM goal is achieved, this leads directly to an increase in cost and time certainty because defects such as those at Berlin's BER Airport are avoided at an early stage [21]. Accordingly, by achieving the BIM goal of increased planning quality, a significant contribution can be made to increasing the degree to which time and cost targets are achieved in factory planning projects.

From this, the two BIM use cases geometric collision checking between partial and functional models and quality checking of the models are derived. In the following, these use cases are assigned to the phases of layout planning and described with the corresponding process.

4.1 BIM use case geometric collision checking

The use case of geometric collision checking between partial and functional models describes the merging of the individual models, their mutual checking for geometric collisions and the subsequent communication of the collisions. For the execution of the use case, it is necessary that models of different disciplines are or at least that several partial models are available in order to be able to check them against each other. Thus, in a little bim approach, this BIM use case is only possible as a check between partial models. [9,11]

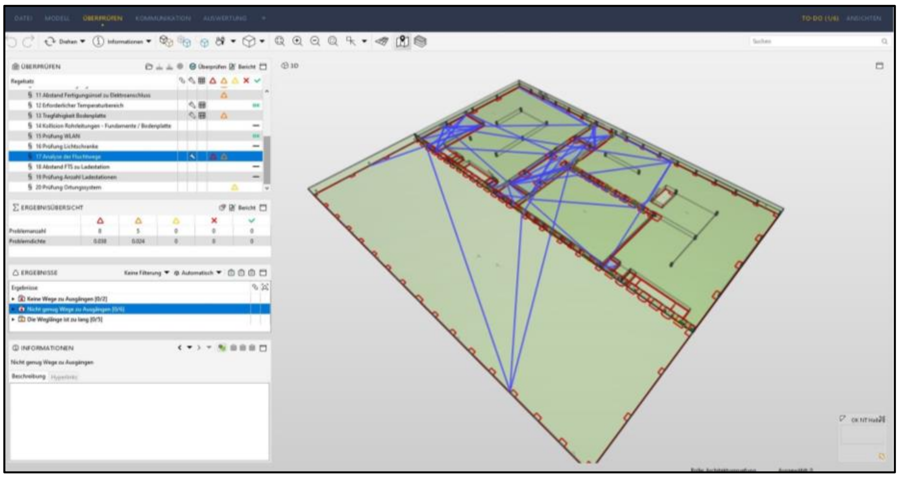
In layout planning, geometric collision checking is usually only useful from the real planning phase onwards, since any collisions in ideal planning are possible but still negligible. This is due to the relatively low LoD of the production system model in the form of a block layout and to the subsequent generation of planning variants, which in turn can differ significantly from the ideal layout. The models should therefore be roughly coordinated with each other, but a software-based collision check does not usually have a positive cost-benefit ratio. [19]

Software-based collision checks are more reasonable in the context of real planning, e.g. whether a production area collides with larger building structures like walls. Detailed collision checks are only useful from the detailed planning phase onwards, as the corresponding LoD is only reached there. For example, it can be checked whether a maintenance and servicing area of a production facility collides with a column.

4.2 BIM use case quality checking

Based on the BIM use case geometric collision check between partial and functional models, the following section focuses on the quality checking of the models in the form of an analysis of the spatial requirements of possible layout variants in terms of their properties in relation to their technical feasibility. For this purpose, it is necessary to design different rule checks to examine restrictions on technical feasibility regarding legal, normative, product-specific or component-specific interfaces and dependencies. [22]

The evaluation criteria for quality assurance are to be analyzed separately in relation to the respective definition of targets, and based on this, corresponding regulatory checks need to be developed. In this context, the examination of legal issues can be identified as an important field of action for the development of quality checks. Accordingly, legal requirements can be identified as first field of action, such as the adherence to the Industrial Building Guideline, e.g. ensuring a maximum escape route length (cf. Figure 4) or the Workplace Guideline, e.g. guaranteeing an appropriate level of illumination in an area of the layout analogous to the work activities to be performed therein. [22]



§ 11 Abstand Fertigungsinsel zu Elektroanschluss
§ 12 Erforderlicher Temperaturbereich
§ 13 Tragfähigkeit Bodenplatte
§ 14 Kollision Rohrleitungen - Fundamente / Bodenplatte
§ 15 Prüfung WLAN
§ 16 Prüfung Lichtschranke
§ 17 Analyse der Fluchtwege
§ 18 Abstand FTS zu Ladestation
§ 19 Prüfung Anzahl Ladestationen
§ 20 Prüfung Ortungssystem

Problemanzahl	8	5	0	0	0
Problemdichte	0.038	0.024	0	0	0

Keine Wege zu Ausgängen (0/2)
Nicht genug Wege zu Ausgängen (0/6)
Die Weglänge ist zu lang (0/5)

Nicht genug Wege zu Ausgängen

Figure 4: Example rule check escape route analysis [22]

In addition, the second field of action, possible product-specific influencing criteria for the case as well as product information, such as material properties, must be analyzed and, if necessary, transferred to the analysis of technical feasibility by means of standard tests. Accordingly, the second field of action describes product-specific influencing parameters, such as manufacturer information on the minimum concrete quality of the fastening substructure for the application of a building product. [22]

Potential component-specific interfaces and, if relevant, existing dependencies between individual components represent a third field of action with potential for the development of rule-based checks. Accordingly, in the third field of action, interfaces and dependencies between individual building components are presented and examined in terms of possible geometric but also qualitative collisions between one another. [22]

For the implementation of the BIM use case, it is necessary, analogous to Chapter 4.1, that models of different disciplines or at least that several partial models are available in order to be able to check them in relation to each other. Thus, in a little bim application, this BIM use case is only possible as a check between partial models and is only possible in layout planning from the real planning phase onwards, since the first non-geometric data is available in the models, which is indispensable for a quality check.

Within the framework of real planning, only conceptual quality checks are useful, e.g. whether the energy requirements of a production area can be covered by the planned energy supply. Detailed checks are only recommended from the detailed planning phase onwards, since again, analogous to Chapter 4.1, the correspondingly required LoD is achieved there. For example, it can be checked whether the load-bearing capacity of the floor is sufficient for a particular piece of equipment or not.

4.3 Quality assurance process

These use cases result in the following BIM process. A BIM coordinator and at least two BIM authors must be available as BIM roles. The BIM coordinator takes over the merging of the functional or partial models, the collision and quality checking as well as the communication of collisions and quality reports and is thus responsible for ensuring quality. The BIM authors create the BIM models and are thus responsible for modeling according to the required LoD. However, before the models are handed over by the authors to the coordinator, it is advisable to first subject their own model to quality assurance. For this purpose, the adherence to the specified designations, the adherence to the agreed model structure, the correct placement within the coordination body, the use of the uniform axis grid, the adherence to the coordinate system as well as the project zero point, the use of the correct model units, a space-filling modeling of the spaces as well as the internal collision freedom should be checked. The two use cases of geometric collision checking and quality assurance should thus be implemented within a functional discipline, as in a little bim approach, before they are applied across disciplines.

BIM authoring software and collision checking software are used as tools for this purpose. A Common Data Environment (CDE) can also be used as a central exchange platform. The data transfer points, i.e. data drops, are to be selected project-specifically and depending on the project progress. In real planning, the models are checked rather sporadically for collisions. In fine planning, on the other hand, the models are merged at regular intervals, e.g. at intervals of one to six weeks. The IFC and BCF formats are used for data exchange in an open BIM approach. The resulting BIM process is shown in Figure 5.

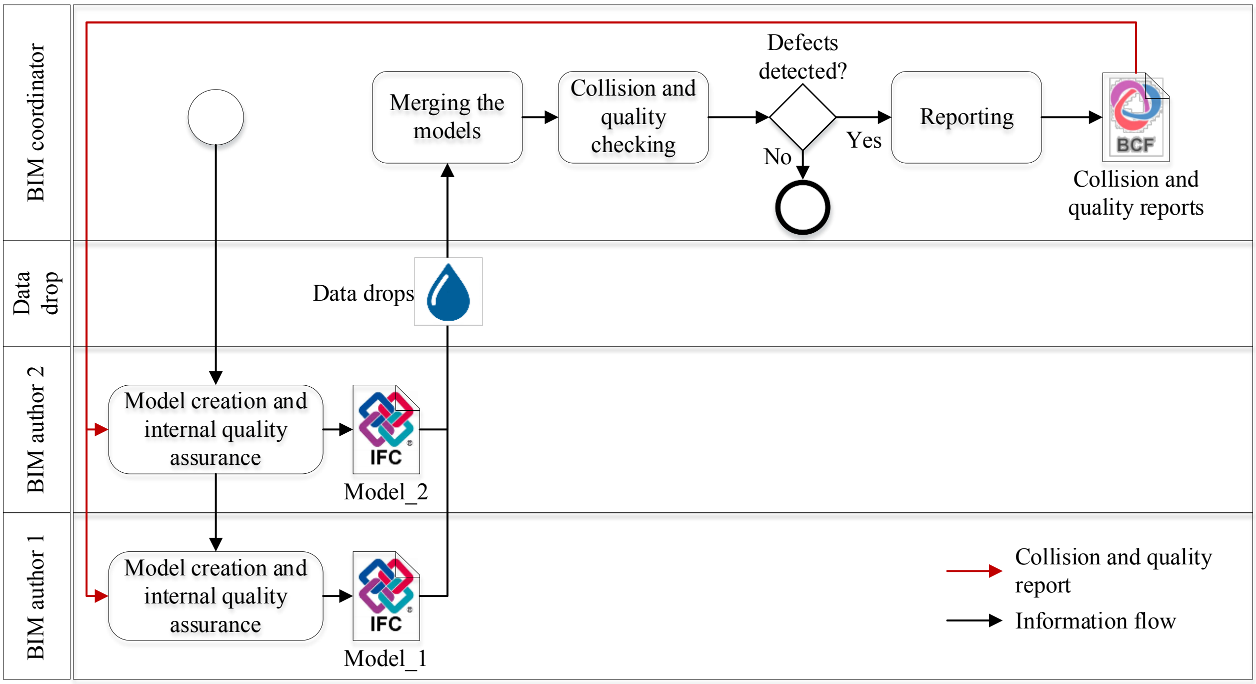


Figure 5: Quality assurance process according to [19]

5. Summary and conclusion

In this paper, a modeling guideline regarding the geometric and non-geometric level of detail in the factory layout planning process, i.e. the Level of Development, is presented. Focusing on the BIM goal of improving planning quality, two BIM use cases, the geometric collision checking of functional and partial models and the quality checking of models are described and merged to a quality assurance process. Thereby, the focus is laid on the layout planning process, since in this phase the production system merges with the building to form the factory and a large part of the planning interaction between the individual trades takes place.

Based on the average of 5,820 completed factory and workshop buildings from 2011 to 2020 in Germany with a total cost volume of 4.7 billion euros and a share of 60% of factory planning projects in which cost targets are missed by about 10%, this results in error costs of about 283 million euros per year [5,23]. The presented potentials of the BIM methodology offer the possibility to avoid these error costs. But how and whether this can also be realized by the approach presented must be evaluated and assessed in the next step on the basis of initial use cases. There are two use cases from the metal processing industry, one Greenfield and one Brownfield project. The Greenfield project is about 25,000 square meters of production and the Brownfield project is about 2,000 square meters. The evaluated and validated results will subsequently be published.

Furthermore, to generate a baseline for BIM-based factory layout planning, more BIM goals and their related use cases will be investigated. Afterwards, this procedure can be extended on the whole factory planning process, as well as the whole factory life cycle. Parts of this follow-up work will be addressed within the framework of *VDI Guideline 2552 part 11.8.1 BIM use case factory layout planning, the VDI expert recommendation BIM-based factory planning* and the *buildingSMART expert group open BIM in factory planning*.

References

[1] Statista, 2021. Anteil der Wirtschaftszweige an der Bruttowertschöpfung in Deutschland im Jahr 2020.

[2] Abele, E., 2011. Zukunft der Produktion: Herausforderungen, Forschungsfelder, Chancen. Hanser, München.

[3] Pawellek, G., 2014. Ganzheitliche Fabrikplanung: Grundlagen, Vorgehensweise, EDV-Unterstützung, 2nd ed. ed. Springer Berlin / Heidelberg, Berlin, Heidelberg.

[4] Wiendahl, H.-P., 2014. Handbuch Fabrikplanung: Konzept, Gestaltung und Umsetzung wandlungsfähiger Produktionsstätten, 2., überarbeitete und erweiterte Auflage ed. Hanser; Ciando, München, Wien, München.

[5] Reinema, C., Pompe, A., Nyhuis, P., 2013. Agiles Projektmanagement. Zeitschrift für wirtschaftlichen Fabrikbetrieb 108 (3), 113–117.

[6] Packwitz, C., Faust, A., 2022. Fehlerkostenbilanz 2020: Rund 18 Milliarden Euro verbauter Schaden. BauInfoConsult. https://bauinfoconsult.de/presse-fehlerkostenbilanz-2020-rund-18-milliarden-euro-verbauter-schaden/. Accessed 16 March 2022.

[7] Borrmann, A., König, M., Koch, C., Beetz, J. (Eds.), 2021. Building Information Modeling: Technologische Grundlagen und industrielle Praxis, 2., aktualisierte Auflage ed. Springer Vieweg, Wiesbaden.

[8] Bundesministerium für Verkehr und digitale Infrastruktur, 2015. Stufenplan Digitales Planen und Bauen: Einführung moderner, IT-gestützter Prozesse und Technologien bei Planung, Bau und Betrieb von Bauwerken.

[9] Meins-Becker, A., Kaufhold, M., 2021. BIM-Handlungsempfehlung: für die kommunalen Bauverwaltungen und die kommunale Gebäudewirtschaft in Nordrhein-Westfalen. Ministerium für Heimat, Kommunales, Bau und Gleichstellung des Landes Nordrhein-Westfalen. www.mhkbg.nrw. Accessed 4 February 2022.

[10]Bryde, D., Broquetas, M., Volm, J.M., 2013. The project benefits of Building Information Modelling (BIM). International Journal of Project Management 31 (7), 971–980.

[11]Tulke, J., Stein-Barthelmes, I., 2021. Umsetzung des Stufenplans „Digitales Planen und Bauen". planen-bauen 4.0 GmbH. https://bim4infra.de/. Accessed 4 February 2022.

[12]Neuhäuser, T., Chen, Q., Rösch, M., Hohmann, A., Reinhart, G., 2020. Building Information Modeling im Fabriklebenszyklus. Zeitschrift für wirtschaftlichen Fabrikbetrieb (s1), 66–69.

[13]Rieke, L., Schäfer, S.F., Hingst, L., Hook, J., Peter, N., 2021. Einsatz von BIM in der Fabrikplanung/Use of BIM in factory planning. wt 111 (11-12), 881–886.

[14]Koeble, W., Zahn, A., 2021. Die neue HOAI 2021: Text und Erläuterungen, 3. Auflage ed. Werner Verlag; Wolters Kluwer Deutschland GmbH, Hürth.

[15]Verein deutscher Ingenieure, VDI-Gesellschaft Produktion und Logistik (GPL), 2011. Fabrikplanung: Planungsvorgehen. Beuth 03.100.99.

[16]Verein deutscher Ingenieure buildingSMART, VDI-Gesellschaft Bauen und Gebäudetechnik (GBG), 2021. Building Information Modeling: Informationsaustauschanforderungen zu BIM-Anwendungsfällen. Beuth 35.240.67.

[17]Grundig, C.-G., 2015. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 5., aktualisierte Aufl. ed. Hanser, München.

[18]Liebsch, P., Sautter, H., 2018. BIM Praxisleitfaden: LOD / LOI Definitionen. Informationen zur Detaillierungs- und Informationstiefe. http://www.bim-blog.de/bim-praxisleitfaden-1-0/. Accessed 4 February 2022.

[19]Neuhäuser, T., Michaeli, P., Lenz, L., Hohmann, A., Matschinsky, F., Madl, F., Henkelmann, R., 2021. Collaborative Factory Layout Planning With Building Information Modeling. Conference on Production Systems and Logistics 2, 699–708.

[20]Hausknecht, K., Liebich, T., 2018. BIM-Kompendium: Building Information Modeling als neue Planungsmethode, 2. überarbeitete und erweiterte Auflage ed. Fraunhofer IRB Verlag, Stuttgart.

[21]Metzner, T., Keilani, F. Dicke Luft wegen zu schwerer BER-Ventilatoren: Statikprobleme in Schönefeld. Tagesspiegel.

[22]Lenz, L., 2020. Bewertungssystem zur Entscheidungsunterstützung von Fabrikgebäudeanpassungen auf Basis von Building Information Modeling. Dissertation.

[23]Statistisches Bundesamt (Destatis), 2021. Baufertigstellungen im Hochbau: Deutschland, Jahre, Bautätigkeiten, Gebäudeart/Bauherr.

Biography

Thomas Neuhäuser (*1990) is group leader for collaborative factory planning at the Fraunhofer Institute for Casting, Composite and Processing Technology IGCV. He is chairman of the VDI-guideline committee BIM-based factory planning and spokesman of the buildingSMART expert group open BIM in factory planning.

Lisa Lenz (*1986) is a civil engineer and works as Post-Doc at the department of construction management at TU Dortmund University. Her research focuses on Building Information Modeling, digitization of the construction industry and research into data-driven rule checks of design services. Dr. Lisa Lenz is managing director of Building Information Cloud GLWG GmbH and spokeswoman of the buildingSMART expert group open BIM in factory planning.

Kai Weist (*1989) is a civil engineer and works as a PhD at the department of construction management at the TU Dortmund University. The main focus of his research lies in the adaptation of Building Information Modeling to factory planning processes and in the digitization of construction processes overall. Since 2020 Kai Weist is managing director and founder of the Building Information Management GLW GmbH.

Andrea Hohmann (*1983) studied aerospace engineering at the University of Stuttgart and received her doctorate in 2019 at the Technical University of Munich on the topic "Life cycle assessment of manufacturing processes for CFRP structures for the identification of optimization potentials". Since 2012, Dr.-Ing. Andrea Hohmann is head of the department sustainable factory planning at the Fraunhofer IGCV.

Rüdiger Daub (*1979) is the holder of the Chair of Production Engineering and Energy Storage Systems at the Technical University of Munich and director of the Fraunhofer IGCV. After studying Electrical Engineering and Information Technology at the Karlsruhe Institute of Technology (KIT), he earned his doctorate at the Technical University of Munich (TUM). From 2012 to 2021, he held several positions at BMW Group Munich, most recently as Head of Technology Development and Prototype Construction of Lithium-Ion Cells.

3rd Conference on Production Systems and Logistics

Application Of Machine Learning On Transport Spot Rate Prediction In The Recycling Industry

Thorben Green, Alexander Rokoss, Kathrin Kramer, Matthias Schmidt

Institute of Product and Process Innovation (PPI) / Leuphana University, Lueneburg, Germany

Abstract

The transport spot rate in trucking logistics is an important factor for market participants in the recycling industry. Knowledge about the current spot rate is essential for operational decision-making in price negotiations between brokers and shippers. Due to the characteristics and dynamics of the industry, this task is particularly challenging. So far, businesses mainly rely on traditional calculation methods combined with their own expertise in price negotiations. The growing amount of existing business and market data may enable companies to take advantage of data-driven decision processes. However, the resulting volume of data and required effort for analysis do not match the fast pace of daily business.

To improve current forecasting practices, this paper conducts a comparative study of machine learning (ML) approaches for shipment-specific spot rate prediction. For this, the paper builds on the experience and database of a small broker in the recycling industry in Northern Germany and complements it with external market information. The study shows the ability of ML to internalize underlying patterns between spot rates and market data. During the use case the CRISP-DM framework is followed to select the most appropriate features and train multiple ML algorithms. Several metrics are applied to determine the most accurate model for spot rate prediction. Results indicate that especially the ML-algorithm Random Forest shows considerable potential to provide brokers in the recycling industry with more reliable spot rate assumptions. Therefore, future implementation of ML approaches in the industry may open up new and beneficial business opportunities. The study paths the way for further research on the predictive potential of ML for prices in transportation with extended and diversified data sets.

Keywords

Machine Learning; Price Prediction; Transport spot rate; Reverse Logistics; Recycling

1. Introduction

The latest supply shortages in Germany and Europe disclosed the urgency to exploit the possibilities of the circular economy for resource recovery to reduce dependency on primary resources [1]. The closing of material cycles by returning waste and product residues to reuse or recycle them is, together with narrowing and slowing the use of materials, at the core of this strategy [2]. Hence, recycling is an essential industry for closed loop supply chains. Along with the collection and storage, the transport of waste products is among the basic processes of logistics in recycling [3]. Since the management of logistics is rarely the core business of the waste owners, specialized companies handle the process. In recycling, a broker arranges the disposal of waste on behalf of others with or without taking possession of it [4]. To their customers they act as a disposer, whereas to carriers and disposal facilities they act as the waste producer. Thereby, they connect

DOI: https://doi.org/10.15488/12151 publish-Ing.

customers and disposal facilities providing efficient and economic transport of waste by contracting an appropriate carrier at the best costs [5]. For shipment, road transport via trucks is often the preferred solution due to their speed, flexibility and versatility [3,6]. Hence, brokers in recycling industry need to seek capacity on the trucking market.

In road transportation, freight rates are distinguished in contract rates and spot rates. Spot rates are a lot more volatile and at times exceed long-term contract rates by more than 60% [7]. Smarter algorithms and sharing economy have led to the rise of digital freight exchange platforms, giving the spot market even more importance [8]. In, addition, the recycling industry is characterized through extensive legal requirements shaping the competition [9]. Further certification and know-how are required to successfully participate in the market, which reflects on less elasticity of the available capacity. Moreover, less-than-truckload shipments are unusual and, depending on the type of waste product, additional permissions and cleaning processes of trucks are mandatory. These aspects enhance the existing uncertainty in price negotiations, as the rates for recycling and transportation of waste products underly high volatility [9]. Therefore, especially for brokers, who qualify through exquisite market knowledge and network, support in price prediction in such a dynamic environment is vital [10].

For spot rate predictions, experience and constant knowledge of market conditions are required. This is becoming a challenge for logisticians due to the ever-increasing amount of data in logistics and the special conditions of the recycling industry. Machine learning (ML) algorithms can be trained with large amounts of data and use this as input to determine a desired response value [11]. It has also been successfully applied to price prediction problems in various industries. The enhanced volatility of spot market prices in the recycling industry and specific infrastructural as well as geographical conditions in Germany amplify the demands on the forecasting model's capabilities. Thus, the study explores the ability of ML models to deal with these settings. The objective of this study is to provide a new prediction approach for market participants that comes with realistic rate suggestions on the spot. The subsequent section reviews the current state of research. Next, ML is applied in a use case based on business data. After identifying relevant features, several ML algorithms are applied to predict future spot rates. The last section summarizes the insights and conclusions for future research.

2. Theoretical background

ML has been applied to price prediction tasks in many research fields (e.g. housing [12], gas [13] and stock prices [14]). In the transport industry, the majority of studies focuses on sea transport (e.g. tanker [15], dry bulk [16] or container freight rates [17]). Studies on price prediction in road transport cover primarily estimation models for spot and contract rates using statistical approaches [18,19]. Some studies also use prediction approaches with ML. Xiao et al. [20] apply GARCH, NN-GARCH and ARIMA models to predict the volatility on the freight rate on the spot market. In another study, Xiao et al. [21] compare a lagged coefficient weighted matrix-based multiple linear regression model with ARIMA and light gradient boosting to predict short term spot rates in southwestern China.

Many studies predict the general price level over time. The forecasts are sometimes further specified to lanes or freight types imparting more specific information. With respect to brokers, simply observing price trends through an estimated index is often insufficient in daily operations. Market players value each transaction based on their individual network and know-how [22]. As every transaction is negotiated individually, spot rates for transport need to be considered shipment-specific. From the broker's perspective, the acceptance of neither the offer to the customer nor the bid to a possible carrier is certain at the time the quote is made to the customer [23]. Therefore, brokers face the challenge of pricing every single transaction with its characteristics in a profitable way to prevent economic risks.

Few researchers have addressed the problem of shipment-specific spot rate prediction. Kay and Warsing [22] develop a non-linear regression model to estimate freight rates for less-than-truckload loads in the US. The model considers various shipment characteristics and public data to provide decision support through shipment price prediction. However, no evaluation of precision of the predicted freight rates is conducted. In a study by Lindsey et al. [24] a data set of a non-asset based broker in the US was used to predict arising carrier costs for the broker. Determinants for carrier costs on spot markets were identified on lane- and shipment-level. In a tactical planning scenario, a regression model was applied to predict carrier costs on unprofitable lanes. The results showed a mean absolute percentage error (MAPE) of 27% on the respective lanes. In another study, Lindsey et al. [10] built a decision making framework for freight brokers on the spot market. The statistical modeling framework consists of decision and profit maximization models. When applied to real-world data, the framework provides a profitable price suggestion for a transaction with potential carriers. Budak et al. [23] applied an artificial neural network and quantile regression to make predictions for spot rates in Turkey in both a route-based and a general model. Route-based predictions emerged more precise than predicting single transactions in the general model. The artificial neural network provides more precise predictions for spot rates in the route-based model (MAPE 0.8 %), while quantile regression performed better in the general model (MAPE 6.7 %).

Few studies focus on analytical predictions on shipment-level and only one applies ML for this task. Therefore, this study aims to build on the shipment-specific studies by Lindsey et al. [24] and Budak et al. [23], by applying a set of ML algorithms for spot rate determination. In addition, the focus of this study on brokers in the recycling industry in Germany leads to special circumstances differing from the general approaches in the literature so far. No other studies covering this part of the circular economy were identified. Therefore, the present study demonstrates the applicability of ML for spot rate determination in the recycling industry by historical and current market data. It determines the best suited algorithms for the prediction task based on a set of evaluation metrics.

3. Methodology

This study follows the widely used Cross Industry Standard Process for Data Mining (CRISP-DM) developed by Chapman et al. [25]. The process contains six phases: “business understanding”, “data understanding”, “data preparation”, “modeling”, “evaluation” and “deployment”. In the business understanding phase, the underlying business objectives, current situation and goals for the project are assessed. During the data preparation phase the final data set used for modeling is built from the initial data. For this, tables and features are selected and additional external data is added. The data is then transformed to be processed during modeling. In the subsequent phase, modeling techniques are determined and applied for the task at hand. This step includes the iterative optimization of the data and model parameters. To determine the quality of the model and to ensure that it meets the expectations of the business, a comprehensive assessment of performance is conducted during the evaluation phase. In the last phase, the final model is deployed. This includes the integration of a closed application supporting or taking over the underlying process in the existing IT-system. The deployment phase is out of scope of this study. Data analysis and modeling is conducted in Jupyter Notebooks operating Python 3.1.0 and using libraries such as Numpy [26], Scikit-learn [27] as well as Pandas [28]. In the subsequent section this methodology will be applied to the use case.

4. Application on the use case

4.1 Business and data understanding

The data for this study is provided by a broker in the recycling industry in Northern Germany. The firm sources transport and recycling capacity for its clients and their waste products. As this study focusses on the road transportation of materials, the price for disposal is out of scope. Since the company is extensively using online freight exchange platforms for sourcing transport capacity, it is exposed to the dynamic rates on the spot market. The business intends to use its data to predict spot rates for future orders. The forecast shall support the order disposition process, where an employee needs to determine the value of an order within a horizon of 1 to 4 weeks based on past data.

The data set for this study consists of several tables containing order, customer, carrier, disposal facilities and product type information. These are merged to a main table consisting 14,244 transactions and 14 variables or features. The target variable in the data set is "freight rate". It is either stored as the total price or as price per ton. However, the spot rates in the transport industry are mainly negotiated as freight rates per km, which are not stored in the initial data set. The transformation of the target variable as well as data inconsistency issues and integration of features are addressed during data preparation.

4.2 Data preparation

The distance and duration for each transaction are obtained through an automated Google Maps API by passing the respective ZIP-codes and locations from the data set. Missing and incorrect data for the target variable are imputed or dropped, resulting in the final, normalized target variable "freight rate per km" (FR/km). The data set then contains categorical and numerical features reflecting temporal, geographical, cargo-specific information. Through an exploratory analysis the impact of features on FR/km is examined. Further data preprocessing steps such as cleaning, grouping, extracting or excluding features and transactions are performed [29]. For example, the lowest and highest 5% of FR/km are removed from the data set, as these rates do not represent the core business activities of the broker.

The numerical features, distance and duration both show a strong inverse correlation (ρ = -0.95), with the target variable. Transactions over short distances are significantly more expensive than long distance trips. Moreover, the time per km ('time/km') can be inferred as an additional feature. Slower trips, for example with longer sections on country roads or through congestion prone areas, show an increase in 'FR/km'. The weight of cargo revealed no interpretable relation to FR/km and was excluded. For categorical features, geographical features are used to group customers and disposal facilities into regions (North, South, East, West or old/new federal state). For trips between the respective regions, differences in FR/km are observed. Product waste types are further summarized into waste classes. As another feature, the carrier of an order is unknown at the point of prediction, unless the order is executed by a broker's truck. Hence, this high cardinality feature is split into two groups (broker, other). Other categorical features are derived from time-related features such as order and transaction date. Information such as day of week, month, quarter or seasons is extracted from them. FR/km shows variation over the course of the year due to differences in capacity availability. Concerning the weekdays, weekend transports are rare and expensive. On Thursdays, market players plan the trips for next week reducing the available capacity and pushing the freight rates. Integrating 'holiday' as an extra feature reveals an increase of FR/km in the days directly before the holiday, especially for short trips.

Some additional, publicly available external features with influence on FR/km are found. The diesel price per liter is available as a time-series data set on the web [30]. Moreover, the truck toll mileage index logs the truck volume on German highways [31]. From this, the available capacity can be inferred. When compared over time, both features accompany the current price level of FR/km and are consequently included in the model. The data preparation procedures result in a final dataset of 11,472 transactions and 23 features.

4.3 Modeling

The prediction task is a regression problem with a continuous target variable. After fitting the model to a training set, it will be tested against unseen data from the hold-out test set [29]. In the order disposition process a lot of past data is used to predict spot rates in a relatively short future horizon (1-4 weeks). However, for modeling, a test set with sufficient data and variety is required. Therefore, the training set is built with 90% or 10,324 transactions from the final data set (from 01.01.2015 to 22.04.2021). The test set consists of 10% or 1,148 transactions from 22.04.2021 to 24.08.2021. This procedure approximates the real application and is referred to as the 90/10 split. Results for the 90/10 split are not generalizable, however, as they could have occurred simply due to the selected composition of the training and test data sets [29]. For more reliable results, cross-validation is applied. In cross-validation the data set is divided into *k*-folds of alternating training and test subsets for each fold. The model is then evaluated by averaging the error of each fold to the overall prediction error [32]. Although the data set is not a time series in particular, the prediction task is time dependent since the goal is to predict exclusively future spot rates from the past. Consequently, conventional cross-validation cannot be used, since it would leak future information that will not be available in a real-world setting. Instead, time-based cross validation using 10 time-dependent folds is applied. By integrating a GridSearch algorithm the ideal parameters for each model are determined.

Before running the model, further preprocessing of numerical and categorical features is required. Numerical features are scaled using the StandardScaler algorithm. Categorical features are encoded using One-Hot Encoding. To reduce loss of performance caused by high dimensionality after encoding, a feature selection process is applied on the data set. Only the most relevant of the encoded features are passed on to the model as boolean features. Several feature selection methods have been compared to determine the number of features leading to the best performance. Out of all compared methods, SelectKBest for regression problems yielded the best results. It was found that performance peaks, when the 11 best features are used for modeling. The final features for the model are listed in Table 1.

Table 1: Final set of selected features for modeling

	Variable	Type	Importance	Meaning
1	curr_diesel_price	numeric	0.31	Current price for Diesel fuel at the time of order
2	distance	numeric	2.10	Distance covered during trip
	distance_cat:	categorical		
3	▪ dist_medium	boolean	0.32	Category, TRUE for trips 60 km to 150 km
4	▪ dist_long	boolean	0.35	Category, TRUE for trips > 150 km
5	truck_toll_index_adj	numeric	0.20	Current Truck Toll Index seasonally adjusted at the time of order
6	frforw_kat_1	boolean	0.18	Category, TRUE for transactions with freight forwarder 1
	direction_transport_old_new_state:	categorical		
7	▪ “new_new“	boolean	0.25	Category, TRUE for trips within new federal states
8	▪ “new_old“	boolean	0.22	Category, TRUE for trips between new and old federal states
	direction_transport_region:	categorical		
9	▪ „north_north“	boolean	0.19	Category, TRUE for trips within regions in Northern Germany
10	▪ „east_east	boolean	0.16	Category, TRUE for trips within regions in East of Germany
11	time_km_breaks	numeric	2.00	Time per km including mandatory breaks [sec]

After feature selection, the authors tested the predictive performance of several ML algorithms on the set of features. As a benchmark, a simple model (BM) is used. The BM reflects estimation principles applied in business so far. During training, the BM calculates individual average spot rates for short (<60km), regional (60-150km) and long (>150km) transports. Surcharges are added for trips directly before public holidays and orders closed on Thursdays, as the data shows significant increases in spot rates in both cases. Regarding the ML algorithm selection, the availability via open source libraries for easy adaptability on this and similar future cases was essential. The study includes simple, more interpretable ML models as well as less intuitive, but possibly more performant, ensemble learning methods that are built from a set of simple base models. Since neural networks are also popular for ML based prediction tasks, a corresponding algorithm was included as well. The simple models Decision Tree (DT) and Lasso Regression (LR) are selected. The ensemble methods consist of the Random Forest algorithm (RF), Gradient Boosting (GB) and eXtreme Gradient Boosting (XGB) [33–35]. Multilayer Perceptron (MLP) is added as a neural-network based algorithm accessible in Scikit-learn.

4.4 Evaluation phase

Evaluation is done using several metrics measuring the prediction error for the target variable of the models for a given order. Performance metrics for regression models have been controversially discussed and employed in literature without finding a common agreement. Still, no consensus on the “best” metric has been achieved [36,37]. To obtain a more reliable evaluation, a combination of metrics should be applied [38]. Therefore, the mean absolute error (MAE), mean absolute percentage error (MAPE), root mean squared error (RMSE) and the coefficient of determination (R^2), are applied as performance metrics in this study. The MAE metric is easily interpretable for practitioners, since it provides a general and bounded performance measure for the model. The MAPE measures relative deviations and is applicable since the target variable contains positive values only. However, MAPE scores are biased towards favoring lower predictions [36]. Even though unrealistic outliers have been removed in the data preparation phase, FR/km shows distinct variance for short distance trips. To assess the ability of models to deal with this variance, RMSE is used as an additional metric. It is more sensitive to variance by giving higher weight to larger errors [38]. R^2 provides a high score, if the greater part of elements is predicted correctly, showing the model’s ability to explain the target variable [39]. Additionally, the computation time for training and prediction is considered as a metric.

5. Results

The results of the modeling phase are evaluated in Table 2. The results of the 90/10 data split are compared to the performance in cross-validation to measure the generalizability of the ML models. The BM was also applied to predict spot rates in the different validation sets of each fold. In general, it can be stated that all ML algorithms outperform the BM in both 90/10 split and cross-validation. In most cases, cross-validation results show a larger MAE, MAPE and RMSE, while R^2 is decreasing.

Table 2: Results of modeling in a 90/10 split and with cross-validation (best three values in bold)

	90/10 Split					**Cross-validation**				
	MAE [€]	RMSE [€]	MAPE [%]	R^2 - Test	Time [sec]	MAE [€]	RMSE [€]	MAPE [%]	R^2 - Test	Time [sec]
DT	**0.109**	0.208	5.97	0.908	2.92	0.142	0.245	8.105	0.745	8.056
LR	0.198	0.301	10.12	0.808	**0.16**	0.176	0.280	9.912	0.680	**7.047**
RF	**0.105**	**0.185**	5.62	**0.927**	9.02	**0.124**	**0.208**	7.275	**0.807**	10.647
GB	**0.106**	**0.194**	5.61	**0.920**	5.21	**0.129**	**0.211**	7.533	**0.801**	**7.895**
XGB	0.115	0.198	6.07	0.917	**0.57**	**0.130**	**0.222**	7.570	**0.787**	**7.234**
MLP	0.129	**0.196**	7.16	**0.918**	7.19	0.179	0.270	10.763	0.663	10.138
BM	0.216	0.324	11.17	0.778	**0.05**	0.206	0.326	11.85	0.642	3.03

In the 90/10 split, RF shows the best performance of all ML algorithms regarding the MAE, with GB and DT being almost equally predictive. However, comparing the RMSE, performance of DT is substantially lower than RF and GB. Moreover, the superiority of RF over GB becomes clearer, when the RMSE is considered. This indicates more stable predictions by RF with less larger errors. LR reveals its limited applicability to non-linear data structures scoring the highest MAE and RMSE and a R² which is only slightly better than the BM model. MLP does not achieve the best scores for the MAE, but presents the third best RMSE. Regarding the computation time, ML algorithms are slower than BM due to the fitting process. XGB and LR achieve the lowest computation time of all ML models.

In cross-validation, the scores for all models worsen. Especially, the DT and MLP algorithms show substantial decrease in performance relatively to their score in the 90/10 split, indicating overfitting tendencies by both algorithms. Again, RF outperforms the other algorithms in most metrics. However, RF presents the worst computing time during cross-validation. XGB and GB show MAE and RMSE results within the range of RF and competitive computation times, making them suitable solutions, when computation time is of importance. Remarkably, LR is the only algorithms that performed better during cross-validation than in the 90/10 split. Yet, it is not competitive in terms of the MAE and RMSE.

Figure 1 shows the distribution of prediction error of the best performing algorithm (RF) as a function of distance. RF predicts especially the dominating long distance trips with high precision. With shorter distances the amount of orders as well as the precision declines. To quantify the potential of using RF, an average transaction in the data set with a distance of 295 km and a mean FR/km of 1.61 €/km is considered. This accounts for average costs of around 475€. Considering the cross-validated MAE of 0.124 €/km, the average deviation is +/- 36.68 € per transport. Using the BM approach, an average deviation of +/- 61.95 € was found. In daily operations, unexpected waiting times are another source for price deviations. For example, a minor delay of 30 min already may cause unplanned, additional charges of around 0.10€/km. Hence, a one-hour delay on an average transaction due to congestion in traffic or at the point of disposal leads to deviations of around + 59.00 €. Therefore, the application of ML, may reduce deviations caused by delays for future transactions.

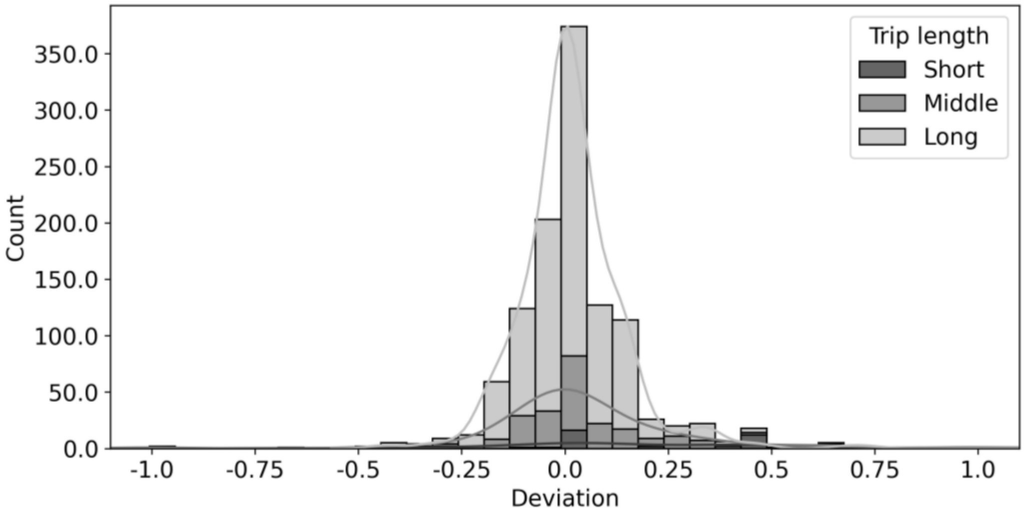


Figure 1: Distribution of prediction error of RF in the 90/10 split

6. Conclusion

The goal of this study was to demonstrate the applicability of ML algorithms for predicting spot rates for transports in the recycling industry. Data from a small broker in Northern Germany was used to train six ML algorithms. The predictive performance was benchmarked against a practical approach. In comparison, ML models outperformed the manually calculated benchmark method, proving the applicability of ML for spot rate prediction. From the set of ML algorithms, the Random Forest regressor minimized the prediction error the most.

The performance of ML on the prediction task in this study confirms promising application opportunities in real-world settings. In fact, performance is likely to improve during operation as more data is gradually fed into the model [32]. Moreover, the prediction horizon will be shorter (1-4 weeks) in comparison to the scenario in the study. More recent data is likely to prove beneficial for the predictive performance of all models. Future research can be dedicated various areas. For example, during modeling, effects of weighting more recent data or excluding older data on model performance can be investigated. Also, more advanced algorithms may yield improved results. Furthermore, the scope of application can be expanded. The effects of a geographical extension on the approach could be explored. Cooperation with other market participants offers opportunities to enlarge the database and improve the basis for modeling. The study not only built a foundation for ML application for spot rate prediction in the recycling industry. Rather, this study sets a starting point for further exploration of ML's predictive potential for price prediction in the transportation industry in research and practice.

Acknowledgements

The study was carried out in the framework of the research project „KI-Werkstatt: Künstliche Intelligenz in Produktionsunternehmen (KIWe)". It is supported by the European Union and the State of Lower Saxony within the EFRE program.

References

[1] Kümmerlen, R., 2021. Zirkuläre Wertschöpfung: Kreislauf bietet Vorteile beim Beschaffen. DVZ Deutsche Verkehrs-Zeitung (46), 2.

[2] Konietzko, J., Bocken, N., Hultink, E.J., 2020. Circular ecosystem innovation: An initial set of principles. Journal of Cleaner Production 253, 119942.

[3] Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K. (Eds.), 2008. Handbuch Logistik, 3. Aufl. 2008 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 1137 pp.

[4] European Union, 2006. Regulation (EC) No 1013/2006 of the European Parliament and of the Council of 14 June 2006 on shipments of waste: Regulation (EC) No 1013/2006.

[5] Johnson, J.C., Schneider, K.C., 1995. Licensed transportation brokers : their joys and frustrations. Transportation journal : a journal of APICS Supply Chain Council 34 (4), 38–51.

[6] Heiserich, O.-E., Helbig, K., Ullmann, W., 2011. Logistik: Eine praxisorientierte Einführung, 4., vollständig überarbeitete und erweiterte Auflage ed. Gabler Verlag, Wiesbaden.

[7] Scott, A., Parker, C., Craighead, C.W., 2017. Service Refusals in Supply Chains: Drivers and Deterrents of Freight Rejection. Transportation Science 51 (4), 1086–1101.

[8] Baron, R., Zintel, M., Zieris, M., Mikulla, D., 2017. Digital platforms in freight transportation: A true industry disruptor? Arthur D. Little.

[9] Kurth, P., Oexle, A., Faulstich, M. (Eds.), 2018. Praxishandbuch der Kreislauf- und Rohstoffwirtschaft. Springer Vieweg, Wiesbaden, 762 pp.

[10] Lindsey, C., Frei, A., Mahmassani, H.S., Park, Y.W., Klabjan, D., Reed, M., Langheim, G., Keating, T., 2014. Predictive Analytics to Improve Pricing and Sourcing in Third-Party Logistics Operations. Transportation Research Record 2410 (1), 123–131.

[11] Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Springer New York, New York, NY.

[12] Park, B., Kwon Bae, J., 2015. Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. Expert Systems with Applications 42 (6), 2928–2934.

[13] Su, M., Zhang, Z., Zhu, Y., Zha, D., Wen, W., 2019. Data Driven Natural Gas Spot Price Prediction Models Using Machine Learning Methods. Energies 12 (9), 1680.

[14] Nikou, M., Mansourfar, G., Bagherzadeh, J., 2019. Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. Intell Sys Acc Fin Mgmt 26 (4), 164–174.

[15] Eslami, P., Jung, K., Lee, D., Tjolleng, A., 2017. Predicting tanker freight rates using parsimonious variables and a hybrid artificial neural network with an adaptive genetic algorithm. Maritime Economics & Logistics 19 (3), 538–550.

[16] Chen, S., Meersman, H., van Voorde, E. de, 2012. Forecasting spot rates at main routes in the dry bulk market. Maritime Economics & Logistics 14 (4), 498–537.

[17] Munim, Z.H., Schramm, H.-J., 2021. Forecasting container freight rates for major trade routes: a comparison of artificial neural networks and conventional models. Maritime Economics & Logistics 23 (2), 310–327.

[18] Miller, J.W., 2019. ARIMA Time Series Models for Full Truckload Transportation Prices. Forecasting 1 (1), 121–134.

[19] Özkaya, E., Keskinocak, P., Roshan Joseph, V., Weight, R., 2010. Estimating and benchmarking Less-than-Truckload market rates. Transportation Research Part E: Logistics and Transportation Review 46 (5), 667–682.

[20] Xiao, W., Gan, M., Liu, H., Liu, X., 2020. Modeling and Prediction of the Volatility of the Freight Rate in the Roadway Freight Market of China. Mathematical Problems in Engineering 2020, 1–15.

[21] Xiao, W., Xu, C., Liu, H., Yang, H., Liu, X., 2020. Short-Term Truckload Spot Rates' Prediction in Consideration of Temporal and Between-Route Correlations. IEEE Access 8, 81173–81189.

[22] Kay, M.G., Warsing, D.P., 2009. Estimating LTL rates using publicly available empirical data. International Journal of Logistics Research and Applications 12 (3), 165–193.

[23] Budak, A., Ustundag, A., Guloglu, B., 2017. A forecasting approach for truckload spot market pricing. Transportation Research Part A: Policy and Practice 97, 55–68.

[24] Lindsey, C., Frei, A., Ali Babai, H., Mahmassani, H.S., Park, Y.-W., Klabjan, D., Reed, M., Langheim, G., Keating, T., 2013. Modeling Carrier Truckload Freight Rates in Spot Markets, in: Transportation Research Board 92nd Annual Meeting (Ed.), TRB 92nd Annual Meeting Compendium of Papers, p. 19.

[25] Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., Wirth, R., 2000. CRISP-DM 1.0: Step-by-step data mining guide.

[26] Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., Del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with NumPy. Nature 585 (7825), 357–362.

[27] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay, 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12 (85), 2825–2830.

[28] McKinney, W., 2010. Data Structures for Statistical Computing in Python, in: Proceedings of the 9th Python in Science Conference. Python in Science Conference, Austin, Texas. June 28 - July 3 2010. SciPy, pp. 56–61.

[29] Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY, 600 pp.

[30] finanzen.net, 2021. Diesel Benzinpreis historische Kurse in Euro | finanzen.net. https://www.finanzen.net/rohstoffe/diesel-benzinpreis/historisch. Accessed 20 December 2021.

[31] Statistisches Bundesamt (Destatis). Truck toll mileage index. https://www.destatis.de/EN/Service/EXDAT/Datensaetze/truck-toll-mileage.html. Accessed 20 December 2021.

[32] Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining: Practical machine learning tools and techniques, Third edition ed. Morgan Kaufmann, Amsterdam, 629 pp.

[33] Breiman, L., 2001. Random Forests. Machine Learning 45 (1), 5–32.

[34] Chen, T., Guestrin, C., 2016. XGBoost, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA. 13 08 2016 17 08 2016. ACM, New York, NY, USA, pp. 785–794.

[35] Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. Ann. Statist. 29 (5).

[36] Armstrong, J., Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting 8 (1), 69–80.

[37] Botchkarev, A., 2019. A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. IJIKM 14, 45–76.

[38] Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. Geosci. Model Dev. 7 (3), 1247–1250.

[39] Chicco, D., Warrens, M.J., Jurman, G., 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ. Computer science 7, e623.

Biography

Thorben Green (*1991) studied industrial engineering at the Leuphana University of Lüneburg. Since 2021, he works as a research associate in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University of Lüneburg. His research focus is on the practical application of machine learning methods in production planning and control.

Alexander Rokoss (*1993) is a research associate and PhD-student in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University Lüneburg since 2018. His research focus is on the practical application of machine learning methods in production planning and control.

Kathrin Julia Kramer (*1991) is an Academic Assistant and PhD-student in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University Lueneburg since 2019. Her research focus is on the practical application of machine learning methods in production planning and control.

Matthias Schmidt (*1978) studied industrial engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he became head of Research and Industry of the IFA and received his habilitation. Since 2018, he holds the chair of production management at the Institute for Product and Process Innovation (PPI) at the Leuphana University of Lüneburg. In addition, he became the head of the PPI in 2019.

3rd Conference on Production Systems and Logistics

Fulfillment of Heterogeneous Customer Delivery Times through Decoupling the Production and Accelerating Production Orders

Tammo Heuer[1], Janine Tatjana Maier[2], Matthias Schmidt[2], Peter Nyhuis[1]
[1]*Institute of Production Systems and Logistics, Leibniz University Hannover, Garbsen, Germany*
[2]*Institute of Product and Process Innovation, Leuphana University of Lüneburg, Lüneburg, Germany*

Abstract

Manufacturing companies are facing increasing customer requirements regarding delivery times and delivery reliability. In this context, customers have different desired delivery times. The fulfillment of heterogeneous customer delivery times represents a major challenge in the competition for customers. If companies succeed in reliably meeting their customers' desired delivery times, this results in an enormous competitive advantage. Instruments for achieving specific delivery times include especially the use of fast-track orders and shifting the customer order decoupling point. When these instruments are used, numerous interdependencies must be considered. Shifting the customer order decoupling point downstream toward a Make-to-Stock production results in higher stock levels. The use of fast-track orders induces longer throughput times for other orders and higher control effort. In this paper, taking these trade-offs into account, an approach is developed that allows delivery time requirements to be met through a systematic determination of the customer order decoupling point and a share of fast-track orders. For this purpose, interdependencies between both instruments and logistic objectives are identified and investigated using logistical models to meet the delivery time requirements at lower logistical costs.

Keywords

Delivery time requirements; Fast-track orders; Rush orders; Customer order decoupling point; Order processing strategy

1. Introduction

In the context of globalisation, companies are facing increasing competition. Products must be manufactured at low cost and high quality. The importance of delivery time and delivery reliability as decisive purchasing criteria continuously increased over time [1,2]. Nowadays, they are considered critical competitive factors [3]. Customers are willing to pay high price premiums to obtain delivery times below standard [4].

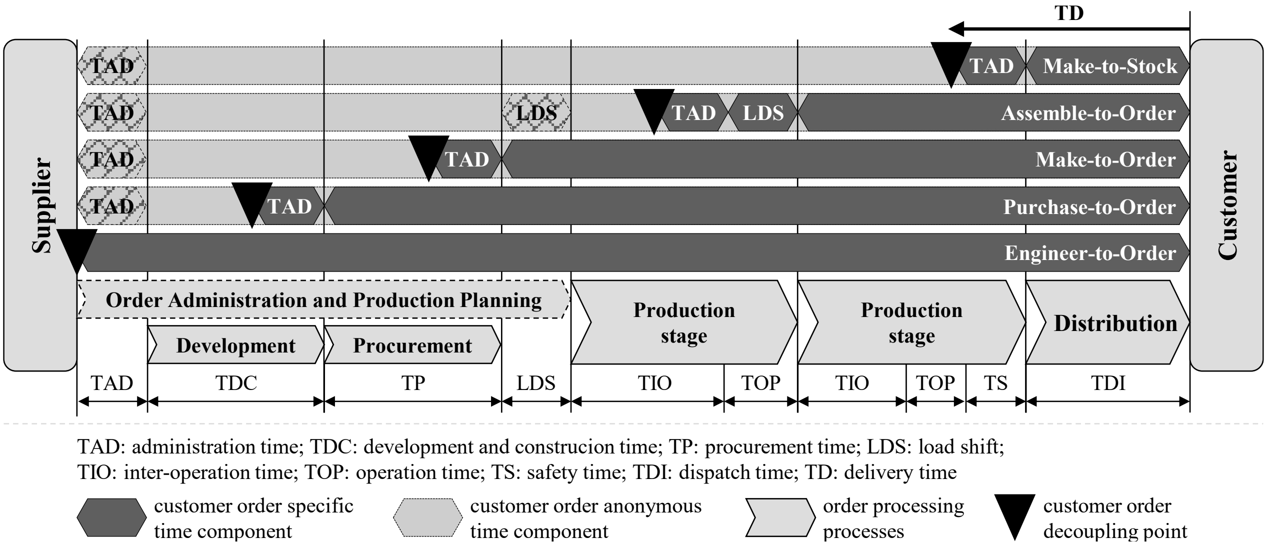
As customer requirements vary widely, companies must face heterogeneous desired delivery times. These can only be mapped entirely by a standard delivery time if the standard delivery time is less than or equal to the minimum requested delivery time. If delivery times below the standard delivery time are requested, these can only be realised by accelerating the orders in the order throughput [5]. Therefore, production planning and control must be designed to handle accelerated orders to avoid undesired effects such as an excessive extension of the throughput times of normal orders. Another central instrument to influence delivery times is positioning the customer order decoupling point (CODP) [4]. However, this can lead to high stocks of finished or semi-finished goods.

DOI: https://doi.org/10.15488/12188

Thus, a holistic view must be taken when designing the production system to map heterogeneous desired delivery times on the production side. As an initial step, section 2 identifies instruments for achieving specific delivery times and describes their suitability. Section 3 provides a literature review on approaches for positioning the CODP and the use of accelerated production orders. In section 4, these essential instruments are examined concerning their interdependencies with logistic objectives. For this purpose, the applicability of existing logistical models is described and possible modeling gaps are identified. Section 5 describes the interactions and trade-offs that need to be considered when using the instruments and setting their parameters. Section 6 summarizes the paper and outlines future research possibilities.

2. Instruments for achieving specific delivery times

There are many instruments to influence the delivery time. They differ in their effects on logistic objectives and their activation time. Mostly they have to be applied long before their effects can be recognized. With most instruments, a short-term reaction to a high number of orders with short delivery times entering the system is impossible. Others can still be used in the short term, although the interactions with the logistic objectives must also be considered. Figure 1 shows the main time components of the delivery time and, depending on the order processing strategy, which time components are effective on the delivery time to the customer.



TAD: administration time; TDC: development and construcion time; TP: procurement time; LDS: load shift;
TIO: inter-operation time; TOP: operation time; TS: safety time; TDI: dispatch time; TD: delivery time

customer order specific time component
customer order anonymous time component
order processing processes
customer order decoupling point

Figure 1: Time components of order processing processes (extension of [6])

According to the order processing strategy, the CODP is positioned differently and time components are either customer order-specific or customer order-anonymous. Thus, the position of the CODP can already be identified as the first elementary instrument on delivery time. There are many different strategies for order processing, but they can be summarised into five main strategies, which are explained below.

In Make-to-Stock (MTS), products are manufactured anonymously in stock, and customer orders are served from these stocks. Assemble-to-Order (ATO) is used when pre-fabricated components or assemblies are assembled to order. Make-to-Order (MTO) describes that production and assembly do not begin until the customer order is received. It is assumed that the raw materials are in stock in all these cases. [7] In Purchase-to-Order, raw materials are also procured on a customer order-specific basis [8]. Lastly, there is Engineer-to-Order, where development and design are customer order-specific [9].

The administration time is required primarily for creating orders [10] and is independent of the location of the CODP. It does not have a significant share in the order throughput but can be reduced, for example, through lean administration approaches [11]. For instance, by using concurrent engineering, development and construction time can be reduced [12]. Procurement time can be influenced through targeted supplier selection or long-term coordination with the supplier. For example, suppliers closer to the production site can be selected or framework contracts can be used to hold safety stocks at the supplier's premises for a short and reliable replenishment time. Additionally, the introduction of consignment stocks from suppliers close to production can reduce the procurement time. [10] In the short term, the procurement time can be shortened through express deliveries or different transport means like air freight [4].

The dispatch time can be influenced in both the short and long term. In the short term, dispatch time can be reduced analogous to the procurement time with express shipping through prioritised treatment in delivery or different transport means like air freight [4]. In the case of MTS production, the dispatch time can be reduced in the long term by setting up a distribution structure. For example, in addition to using the factory warehouse solely, stocks can also be stored close to the customer in central or regional warehouses to guarantee shorter delivery times [13]. However, this requires an accurate forecast of the customer demand.

Within the production stages, the load shift, the inter-operation time, the operating time and the safety time can be influenced. In all order processing strategies with an order-specific production run, an extension of the delivery time can be caused by a load shift. A load shift occurs when capacities are already utilised for a more extended period and new orders can only be scheduled later on when capacities are free [4]. Within this framework, the provision of capacity flexibility is a possibility to accept new orders and realise short delivery times by using additional capacity at short notice [14]. To this end, measures must be taken to create capacity flexibility. However, capacity flexibility is limited and cost-intensive [14]. Capacity can also be reserved within the framework of production planning to be able to schedule orders at short notice [15]. But if short-term orders fail to materialise, capacity utilization losses are a risk. To avoid this, the capacity reservation can be combined with a work-in-process (WIP) regulating order release so that orders are brought forward. [4] On the one hand, this leads to a negative schedule deviation and thus to finished goods stocks due to premature order completion. On the other hand, short-term orders can be realised without using capacity flexibility and without overloading the production system, which would result in a backlog.

The safety time can be scheduled as part of throughput scheduling to compensate for possible delays from production and still deliver on time. In this way, a safety time increases the delivery reliability, but simultaneously increases finished goods stock due to premature order completion. [16] To reduce the delivery time at short notice, less safety time or no safety time can be assigned during throughput scheduling. This might decrease the delivery reliability. The operation time can be reduced at short notice by splitting lots, assuming the availability of work systems that can be used in parallel. Work processes can also be carried out overlapping. [17] In the medium term, transport processes between work systems can be optimised. This aims at reducing the minimum transition time, which is the transport time as part of the inter-operation time [2]. A central possibility for shortening the inter-operation time is the use of accelerated orders [17]. Rush orders can be used to achieve maximum speed-up for time-urgent orders by prioritising these orders to the front of a queue [5]. Fast-track orders differ from rush orders as they are scheduled with individual inter-operation times just to meet the delivery date and are sequenced by the due date. Fast-track orders can thus also be accelerated as much as rush orders. However, this is only done when necessary. As a result, normal orders will probably not be delayed as much and higher shares of fast-track orders can be accepted than is the case with rush orders. [18]

It can be concluded that the positioning of the CODP and the use of accelerated orders are easy means to influence the delivery time in a targeted manner in comparison with the other means introduced before. Therefore, in section 3, existing approaches of achieving specific delivery times through fast-track orders and shifting the CODP are discussed to highlight the need for further research.

3. Literature review and need for research

TRZYNA modeled the throughput time of rush orders and, taking into account a rush order share, the throughput time of normal orders for single work systems [19]. This enables the calculation of the minimal achievable throughput times for given production systems. LÖDDING AND ENGEHAUSEN have developed an approach to maximise incoming orders in the context of heterogeneous desired delivery times using rush orders. This approach is based on pure MTO production. Therefore, only the rush order share is used as a control variable to realise different delivery times. In addition, simplifying a single-stage production is assumed. [20] Due to these assumptions, mainly the focus on MTO, the approach is limited for the application under consideration. Many other approaches focus on achieving different delivery times by using rush orders in an MTO production by providing production planning procedures. CHUNG ET AL. describe an order release that releases rush orders in a stock-controlled manner or at the same time interval so that the production system is not overloaded [21]. Some approaches focus on rescheduling when rush orders occur [22–24]. LIU AND LIU consider production and distribution together and minimise the delivery time using linear optimisation for scheduling, batching and delivering [25]. These works are essential for improving the usage of accelerated orders. However, they may not support the choice of whether to prefer accelerated orders or a shift of the CODP. Regarding this choice, the research done so far only provides for an upper limit of approx. 30% of the work content for rush orders. Otherwise, the rush orders would interfere with each other resulting in an increase in their mean throughput times and the throughput time variation [26,15].

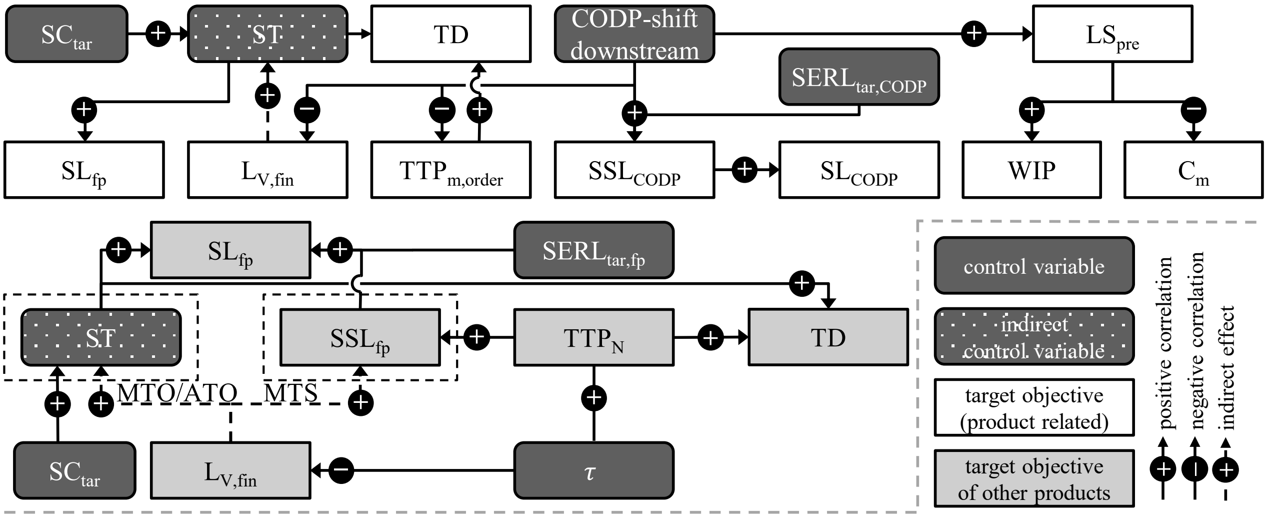
A systematic description of influencing the delivery time through shifting the CODP was performed by HOEKSTRA AND ROMME [8]. TEIMOURY ET AL. determine the appropriate position of the CODP using a linear optimisation model taking into account the product costs [27]. Most authors provide a procedure in which the CODP is positioned to achieve the lowest required delivery time. GRIGUTSCH developed a model-based positioning of the CODP. Contrary to other authors, this author clearly addresses delivery time and schedule compliance. [28] In the research project on which this paper is based, an approach was developed by MAIER ET AL. linking logistical models. This approach makes it possible to calculate which order processing strategy has the lowest logistical costs for each product, achieving a certain schedule compliance or a certain service level in the finished-goods store. In doing so, fixed delivery time requirements per product are assumed to narrow down the solution space. [29]

In summary, there is no approach yet that enables companies to design their production through the choice of order processing strategy and the use of fast-track orders so that heterogeneous desired delivery times can be served, but low logistical costs are achieved. Approaches to the use of fast-track orders are usually only geared to their processing within the framework of MTO production. Approaches to the choice of order processing strategy usually consider a fixed delivery time to be achieved and neglect the possibility of producing orders with a lower than the standard planned throughput time. For example, this can result in MTS being planned for a product to guarantee the fixed delivery time, resulting in high costs due to finished goods stock. However, the proportion of orders for this product that require a short throughput time could also be handled in the context of MTO production through fast-track orders. That would mean that no stocks of finished goods would have to be kept. However, interactions of the fast-track orders with other orders have to be taken into account.

Therefore, this paper aims to systematise these interdependencies to select the order processing strategy that would result in the lowest logistical costs under the assumption of the same logistical performance in service level (MTS) or schedule compliance (MTO, ATO) when using fast-track orders.

4. Effects of using fast-track orders and shifting the CODP on logistic objectives

Different order processing strategies with different needs for accelerating production orders are possible to serve heterogeneous desired delivery times. Therefore, this section describes the interdependencies that need to be considered when deciding between these options. The effects of shifting the CODP and using fast-track orders on logistic objectives and interactions were summarised in Figure 2 and are explained in the following.



C_m	manufacturing costs [€ / unit]	SL_{fp}	finished products stock level [units]
$L_{V,fin}$	variation of lateness in the final production stage [hrs.]	SSL_{fp}	finished products safety stock level [units]
LS_{pre}	lot sizes in pre-production stage [units]	ST	safety time [SCD]
SCD	shop calendar day	TD	delivery time [days]
SC_{tar}	target schedule compliance [%]	$TTP_{m,order}$	order specific mean throughput time [SCD]
$SERL_{tar,CODP}$	target service level in CODP [%]	TTP_N	throughput time of normal orders [SCD]
$SERL_{tar,fp}$	target service level of finished products [%]	WIP	work in process (WIP) [hrs.]
SL_{CODP}	CODP stock [units]	τ	share of fast-track orders on the work content [%]
SSL_{CODP}	CODP safety stock level [units]		

Figure 2: Effects of fast-track orders and position of the CODP on logistic objectives

A downstream shifted CODP (see the upper part of Figure 2) allows larger, more economical lot sizes LS_{pre} in the pre-production stage. Lot size calculation rules can be taken from [30]. Larger lot sizes reduce the manufacturing costs C_m due to lower set-up costs. However, larger lot sizes also affect a higher level of WIP in the pre-production stage. Furthermore, the shift affects a higher CODP stock of semi-finished products SL_{CODP}. [31] The calculation of the safety stock level SSL_{CODP} and the stock level SL_{CODP} depending on the target service level $SERL_{tar,CODP}$ can be taken from [32]. Since a smaller part of the order throughput is customer order-specific (see Figure 1), the throughput time $TTP_{m,order}$ for the product in question decreases. In addition, the variation of the lateness in the final stage $L_{V,fin}$ is reduced for this product as fewer work systems are passed through [28]. Thus, shorter safety times ST can be used to achieve the target schedule compliance SC_{tar}. As a result, the delivery time TD and the finished products stock level SL_{fp} due to orders completed too early are reduced [16].

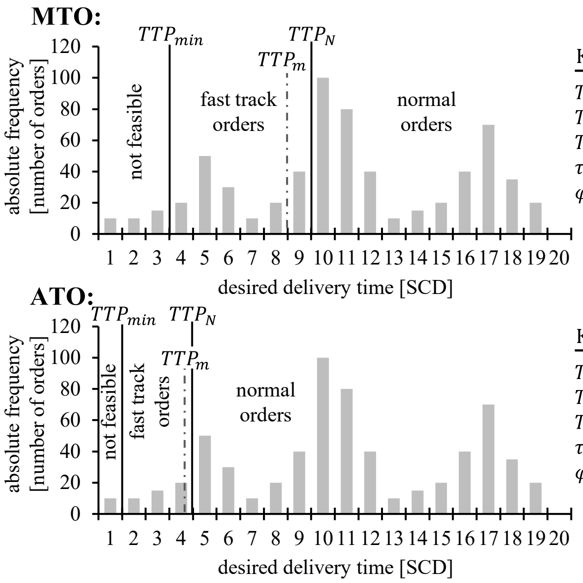
Shorter delivery times can also be realised by integrating fast-track orders into production [20]. However, the integration of a share of fast-track orders on the work content τ has the consequence that the throughput times of normal orders TTP_N are extended [18,19]. Similarly, the variation of the lateness of normal orders in the final production stage $L_{V,fin}$ may increase [18]. This interaction still has to be modeled to enable a holistic model-based evaluation of occurring logistical costs. Higher variation of lateness $L_{V,fin}$ and the longer throughput time of normal orders TTP_N result in higher safety stocks SSL_{fp} for MTS products. The calculation of SSL_{fp} and SL_{fp} in accordance with $SERL_{tar,fp}$ can also be taken from [32]. In case of MTO or ATO products, with a higher variation of lateness $L_{V,fin}$ higher safety times ST have to be allocated to reach the target schedule compliance SC_{tar}. How to dimension the safety time ST in accordance with the variation of the lateness $L_{V,fin}$ and the target schedule compliance SC_{tar} and which delivery time TD and which finished product stock level SL_{fp} result from this safety time, can be determined following [16].

5. Analysis of the delivery time related suitable order processing strategy

Based on the previously identified interdependencies between the positioning of the CODP and the use of fast-track orders with logistic objectives, this section examines interactions and trade-offs between the instruments and their parameterisation.

It is assumed that delivery times below the mean throughput time are served with fast-track orders if it is possible for the considered CODP position. Therefore, the position of the CODP determines how many orders can be served at all and how many fast-track orders must be used with which inter-operation time reduction. A product view and a resource view are described below for analysis purposes.

The product view is shown in Figure 3 with a fictional example for producing a single product. The diagrams show the absolute frequency of delivery times requested by customers in the number of orders in a reference period. The delivery times considered here have been adjusted for time components such as dispatch time so that it is essentially a maximum permissible throughput time. For simplicity, the term delivery time will continue to be used in the following. Three possible scenarios are MTS, ATO and MTO. Since there is no order-related production throughput time for MTS, the logistical costs for MTS with a specific target service level can be calculated following the interdependencies from section 4. Therefore, only MTO and ATO are considered concerning the fulfillment of the heterogeneous desired delivery times.



Key figures

TTP_{min} = 4 SCD
TTP_N = 10 SCD
TTP_m = 9 SCD
τ = 27 %
φ = 6 %

Caption

SCD	shop calendar day
TTP_{min}	minimal througput time [SCD]
TTP_N	throughput time of normal orders [SCD]
TTP_m	mean throughput time of all orders [SCD]
τ	share of fast track orders reg. this product (group) [%]
φ	share of rejected orders [%]

Key figures

TTP_{min} = 2,00 SCD
TTP_N = 5,00 SCD
TTP_m = 4,87 SCD
τ = 7 %
φ = 2 %

Figure 3: Fulfilling heterogenous desired delivery times with different order types

It is assumed that a rush order's progressing starts at each work system as soon as the last order in progress finishes (see [18]). All desired delivery times below this cannot be served by accelerated orders either, which leads to a rejection rate φ. In this example, the logistical positioning of planned WIP at MTO results in a mean throughput time TTP_m of 9 shop calendar days. Fast-track orders achieve delivery times shorter than the mean throughput time. The equilibrium condition, according to TRZYNA and HEUER ET AL., provides an approach to determine the planned throughput time for normal orders TTP_N as the boundary between normal and fast-track orders using the distribution of the desired delivery times, the minimum realisable throughput time TTP_{min} and the mean throughput time TTP_m (see Figure 4) [18,19]. All orders with a delivery time higher than the minimum delivery time and less than the planned throughput time for normal orders are therefore fast-track orders, which in their entirety give rise to a fast-track order share of τ in all orders.

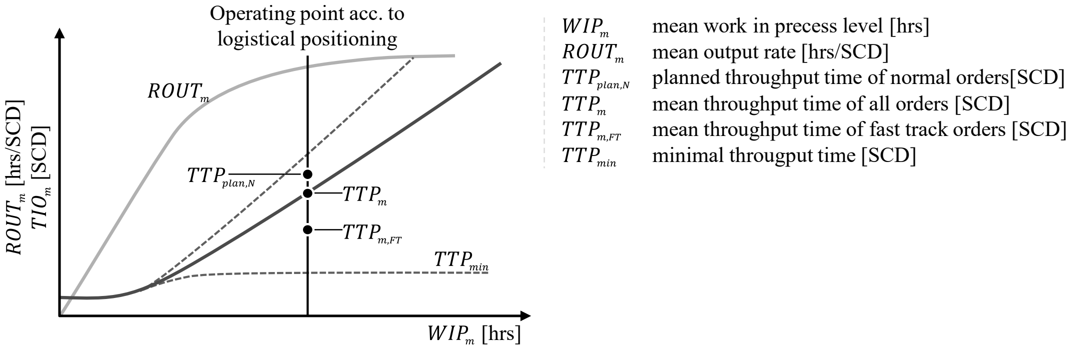


Figure 4: Output and Throughput Time Operating Curves with fast-track orders [18]

In the example for MTO, φ is 6%. These potentially lucrative orders cannot be accepted. With ATO, on the other hand, only 2% cannot be accepted. While for MTO both production stages are burdened with 170 fast-track orders, for ATO only the final production stage is burdened with 35 fast-track orders.

Due to the interactions with the logistic objectives of other products described in section 4, the product view described must be expanded to include a resource view to identify possible problems at the production stages under consideration and thus also other products (see Figure 5).

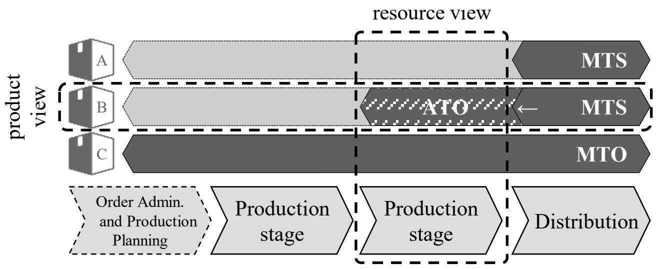


Figure 5: Product view and resource view

When deciding on a shift of the CODP or the use of fast-track orders, the same decision must always be made for other products. If the CODP for a product is set so that fast-track orders become necessary at a production stage, this must be taken into account in the resource view. The desired delivery times must be broken down to the maximum permissible throughput times in the order-specific production run on the individual production stages to be run through. Special effects such as a limitation of the inter-operation time reduction at work systems with set-up time-optimising sequencing may have to be considered. The resource view of a work system must be aggregated for all products whose maximum permissible throughput times are to be processed across all products. In this way, the resource view can derive planned throughput times for standard orders.

For example, a production with two MTS products (product A and B) and one MTO product (product C) is assumed. Product B is switched from MTS to ATO to save costs (see Figure 5). To achieve the desired delivery time of the customer without shifting the COPD downstream, orders of product B in the final production stage have to be partially scheduled as fast-track orders. To check the profitability of this change, the interdependencies shown in Figure 2 must be taken into account. By scheduling fast-track orders in the final production stage, the throughput times of all other orders in this production stage are extended. This means that the throughput time of products A and C are extended. Thus, with a higher replenishment time, the safety stock of product A in the finished goods stock must be increased to achieve the same service level for product A. While an increase in safety stock can compensate for longer replenishment times for product A, it must be checked for product C whether the desired delivery times can still be realised despite this throughput time extension. If this is critical at the final production stage, there is still the possibility of using acceleration potentials at the production pre-stage for product C. To assess the profitability of changing the order processing strategy of product B, the initial savings in the stock of product B must be weighed against the cost of the additional demand for safety stock for product A and secondary effects of product C.

6. Conclusions

Meeting customers' heterogeneous delivery time requirements holds great potential for companies to increase customer satisfaction and revenues. There are many instruments to realise shorter delivery times. Some of them should be chosen strategically and in the long term, others can be used in the short term. As two key instruments, the positioning of the CODP and the use of fast-track orders were investigated in this paper. While the positioning of the CODP, in particular, reduces the average throughput time, a specific range of different delivery times can be mapped through the systematic use of fast-track orders. So far, however, there is no approach that takes into account a coupled decision on the position of the CODP and the use of accelerated production orders, potentially resulting in higher logistical costs.

When positioning the CODP, numerous interactions with logistic objectives have to be considered. The positioning of the CODP per product cannot be done in isolation, assuming the use of fast-track orders. In addition, interactions with other products must be taken into account from a resource perspective. This paper describes the most important influences on logistic objectives and the interactions on the resource view. That makes the coupled positioning of the CODP using fast-track orders accessible to scientific research. Thus, a first approach for determining resource-based planned throughput times could already be developed to link the decision on product level and its influences on other products. That already allows a trade-off in the product-related decision to shift the CODP. The use of heterogeneous planned throughput times requires precise design but can meet heterogeneous delivery time requirements at the same target utilisation rates and lower stocks, thus being more competitive.

The approach proposed in this paper for the positioning of the CODP in the context of the use of fast-track orders must be concretised. Modeling gaps, like the interaction of the share of fast-track orders on the variation of lateness, have to be quantified. As a result, it will be possible to determine for the entire product portfolio which position of the CODP should be selected, considerung fast-track orders, in order to achieve minimum logistical costs in the context of heterogeneous delivery time requirements for a given target logistical performance.

Acknowledgements

The research project was carried out in the framework of the industrial collective research programme (IGF no. 20906 N). It was supported by the Federal Ministry for Economic Affairs and Climate Action (BMWK) through the AiF (German Federation of Industrial Research Associations eV) and the BVL (Bundesvereinigung Logistik eV) based on a decision taken by the German Bundestag.

References

[1] Handfield, R.B., Straube, F., Pfohl, H.-C., Wieland, A., 2013. Embracing global logistics complexity to drive market advantage. DVV Media Group, Hamburg.
[2] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2015. Handbook factory planning and design. Springer, Berlin, Heidelberg, 501 pp.
[3] Gudehus, T., Kotzab, H., 2012. Comprehensive logistics, 2., rev. and enl. ed. Springer, Berlin, Heidelberg.
[4] Lödding, H., 2013. Handbook of Manufacturing Control: Fundamentals, description, configuration. Springer, Berlin, Heidelberg.

[5] Trzyna, D., Kuyumcu, A., Lödding, H., 2012. Throughput time characteristics of rush orders and their impact on standard orders. Procedia CIRP (3), 311–316.

[6] Sharman, G., 1984. The rediscovery of logistics. Harvard Business Review 62 (5), 71–80.

[7] Wemmerlöv, U., 1984. Assemble-to-order manufacturing: Implications for materials management. Journal of Operations Management 4 (4), 347–368.

[8] Hoekstra, S., Romme, J., Argelo, S.M., 1992. Integral Logistic Structures: Developing Cus-tomer-Oriented Goods Flow. Industrial Press, New York.

[9] Wortmann, J.C., 1983. A Classification Scheme for Master Production Scheduling, in: Wilson, B., Berg, C.C., French, D. (Eds.), Efficiency of Manufacturing Systems. Springer US, Boston, MA, pp. 101–109.

[10] Schönsleben, P., 2012. Integral Logistics Management: Operations and Supply Chain Management Within and Across Companies, 5th ed. Chapman & Hall/CRC Press, Boca Raton.

[11] Chen, J., Cox, R., 2012. Value Stream Management for Lean Office—A Case Study. American Journal of Industrial and Business Management 2 (2), 17–29.

[12] Sohlenius, G., 1992. Concurrent Engineering. CIRP Annals 41 (2), 645–655.

[13] Chopra, S., 2003. Designing the distribution network in a supply chain. Transportation Research Part E: Logistics and Transportation Review 39 (2), 123–140.

[14] Kingsman, B.G., Tatsiopoulos, I.P., Hendry, L.C., 1989. A structural methodology for managing manufacturing lead times in make-to-order companies. European journal of operational research 40, 196–209.

[15] Thürer, M., Silva, C., Stevenson, M., 2010. Workload control release mechanisms: from practice back to theory building. International journal of production research 48 (12), 3593–3617.

[16] Schmidt, M., Bertsch, S., Nyhuis, P., 2014. Schedule compliance operating curves and their application in designing the supply chain of a metal producer. Production Planning & Control 25 (2), 123–133.

[17] Mertens, P., 2013. Integrierte Informationsverarbeitung 1: Operative Systeme in der Industrie, 18th ed. Springer Fachmedien Wiesbaden, Wiesbaden.

[18] Heuer, T., Maier, J.T., Schmidt, M., Nyhuis, P., 2021. Be in control of rush orders logistically. wt Werkstattstechnik online 111 (4), 185–189.

[19] Trzyna, D., Lödding, H., Nyhuis, P., 2015. Modellierung und Steuerung von Eilaufträgen in der Produktion. Zugl.: Diss. TUHH Institut für Produktionsmanagement und -technik, Hamburg.

[20] Lödding, H., Engehausen, F., 2019. Use Rush Orders Strategically: How Make-to-Order Companies Can Increase Their Order Intake. Zeitschrift für wirtschaftlichen Fabrikbetrieb 114 (7-8), 449–454.

[21] Chung, Y.H., Seo, J.C., Kim, C.M., Kim, B.H., Park, S.C., 2017. Reservation-based dispatching rule for make-to-order wafer FAB with high-priority lots. Concurrent Engineering 25 (1), 68–80.

[22] He, X., Dong, S., Zhao, N., 2020. Research on rush order insertion rescheduling problem under hybrid flow shop based on NSGA-III. International journal of production research 58 (4), 1161–1177.

[23] Psarommatis, F., Zheng, X., Kiritsis, D., 2021. A two-layer criteria evaluation approach for re-scheduling efficiently semi-automated assembly lines with high number of rush orders. Procedia CIRP 97, 172–177.

[24] Ren, X., Wang, X., Geng, N., Jiang, Z., 2021. The Just-In-Time Job-Shop Rescheduling with Rush Orders by Using a Meta-Heuristic Algorithm, in: 17th International Conference on Automation Science and Engineering (CASE), Lyon, France. IEEE, pp. 298–303.

[25] Liu, L., Liu, S., 2020. Integrated Production and Distribution Problem of Perishable Products with a Minimum Total Order Weighted Delivery Time. Mathematics 8 (2), 146.

[26] Jäger, Y., Roser, C., 2018. Effect of Prioritization on the Waiting Time, in: Moon I., Lee G., Park J., Kiritsis D., von Cieminski G. (Ed.), Advances in Production Management Systems. Production

Management for Data-Driven, Intelligent, Collaborative, and Sustainable Manufacturing, vol. 535. Springer, Cham, pp. 21–26.
[27] Teimoury, E., Modarres, M., Khondabi, I.G., Fathi, M., 2012. A queuing approach for making decisions about order penetration point in multiechelon supply chains. Int J Adv Manuf Technol 63 (1-4), 359–371.
[28] Grigutsch, M., 2016. Modellbasierte Bewertung der logistischen Leistungsfähigkeit in Abhängigkeit des Kundenauftragsentkopplungspunktes. Dissertation.
[29] Maier, J.T., Heuer, T., Stoffersen, H., Nyhuis, P., Schmidt, M., 2022. Data based analysis of order processing strategies to support the positioning between conflicting economic and logistic objectives. Procedia CIRP, (Full Paper accepted).
[30] Münzberg, B., 2013. Multikriterielle Losgrößenbildung. PZH-Verlag, Garbsen.
[31] Nywlt, J., 2016. Logistikorientierte Positionierung des Kundenauftragsentkoppelungspunktes. Dissertation.
[32] Nyhuis, P., 2009. Fundamentals of production logistics: Theory, tools and applications, Softcover reprint of the hardcover 1st edition 2009 ed. Springer, Berlin, Heidelberg.

Biography

Tammo Heuer (*1992) studied industrial engineering at the Leibniz University Hannover and has been working as a research associate at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hannover in the field of production management since 2018.

Janine Tatjana Maier (*1994) studied industrial engineering at the Leibniz University Hannover. Since 2018, she works as a research associate in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University of Lüneburg.

Matthias Schmidt (*1978) studied industrial engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he became head of Research and Industry of the IFA and received his habilitation. Since 2018, he holds the chair of production management at the Institute for Product and Process Innovation (PPI) at the Leuphana University of Lüneburg. In addition, he became the head of the PPI in 2019.

Peter Nyhuis (*1957) studied mechanical engineering at Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he received his habilitation before working as a manager in the field of supply chain management in the electronics and mechanical engineering industry. He is heading the IFA since 2003. In 2008 he became managing partner of the IPH - Institut für Integrierte Produktion Hannover gGmbH.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2022

3rd Conference on Production Systems and Logistics

Business Model Scenarios For Digital Textile Microfactories

Marcus Winkler[1], Franziska Moltenbrey[1], Meike Tilebein[2,1]
[1]German Institutes of Textile and Fiber Research, Denkendorf, Germany
[2]University of Stuttgart, Institute for Diversity Studies in Engineering, Stuttgart, Germany

Abstract

In several industries the concept of Microfactories has been developed and their potentials are still subject of research and development. Also in the Textile and Clothing Industry this concept of a digitally networked end-to-end digital design and production process finds its first realizations in different applications. Such a Digital Textile Microfactory can cover a complete value creating chain comprising all design and production steps from the customer to the ready-made product. It relies on virtual models of the process steps involved, as well as of the materials, the products, and the customers. The digital backbone allows for speed, efficiency, high quality, and deep consumer interaction leading to a great innovation potential in a wide area of applications and Business Models. They range from B2B types, where they can support and speed up the prototyping phase of product development (sampling) up to B2C settings for the innovative production of individualized products, including reordering as well as event-driven production and locally centred production. But in spite of the potential benefits of a Digital Textile Microfactory, there are still just a few realizations seen in the industry due to investment risks and uncertainty with regards to new Business Models.

The goal of this paper is to explore Business Model scenarios for a Digital Textile Microfactory that uses digital textile printing as a core process. We first describe the economic characteristics of the Textile and Clothing Industry, and then the digital technology and process underlying such a Digital Textile Microfactory. Based on this description, we then explore different B2B and B2C application scenarios – developed in previous European and German research projects – and settings for related Business Models.

Keywords

Microfactory; Digital Textile Printing; Business Model; Textile Industry; Digitalization

1. Introduction

The Textile and Clothing Industry (TCI), and especially the sectors of fashion and apparel as well as home and household textiles are characterised by global value chains and by high dynamics. Reasons for this include the megatrend individualization [1], worldwide competition, as well as recent technology and process innovations [2]. The traditional reactions of cost reductions, outsourcing of production to low-wage countries or unspecific stimulation of product demand trying to restore mass production are not promising for many new Business Models, and thus do not always offer an adequate answer to the extreme short product life cycles for many companies in the fashion sector.

Specific additional challenges in the fashion sector arise from the facts that first, product development usually takes relatively long time, as it involves manual process steps and iterated production of physical

DOI: https://doi.org/10.15488/12141

samples and second, end-consumer demand is highly volatile, resulting in stock-outs or wasteful overproduction.

Moreover, apart from these economic characteristics the TCI is faced with social and ecological grievances, such as child labour and environmental pollution through chemicals playing a role in end consumers' purchasing decisions. This can be seen in approaches such as Circular Fashion, Sharing Economy [3] and the consideration of sustainability aspects, e.g. ecological footprint, use of resources, cradle-to-cradle approaches or the large number of different labels and certificates, signalling sustainability aspects towards customers.

Digital technologies deriving from the vision of Industry 4.0 offer great potential to address these challenges for the TCI. Solutions for value creation networks and production using intelligent networking [5], taking into account the possibilities of flexible production, adaptable factories, customer-centric solutions, optimised logistics, use of data for new Business Models and resource-saving circular economy [6], that have proven to be beneficial to other industries can also be applied to the TCI.

Among these approaches Microfactories have recently been subject to research and development [7] and offer new possibilities for the TCI in the form of a Digital Textile Microfactory (DTMF). A DTMF is an end-to-end digitally networked development and production process for textile and clothing products. Its digital backbone allows for speed, efficiency, high quality, and deep consumer interaction leading to a great innovation potential in a wide area of applications and Business Models. A DTMF can provide solutions for efficient development and production of individualised products in small lot sizes, being fast and flexible in a more sustainable production and supply chain.

Consistent digitalization along the entire value chain and the associated information transparency as included in the DTMF also offer solutions to issues of ecological and social sustainability.

Within the next sections of this paper the basic principles and process steps of a DTMF will be presented (section 2), followed by a description of different B2B and B2C Business Model scenarios for the TCI. In particular, these Business Models address firstly the production of individualised products, the sampling of prototypes, the re-ordering (of best-sellers) and the event-driven and locally centred production by using different kinds of settings of the DTMF. They will be discussed in depth in section 3 and summarized in section 4.

2. The Digital Textile Microfactory

The first forms of Microfactories were described in Japan in the 1990s [8]. The DTMF has already existed as a technical implementation in different forms and settings for several years. The Adidas Group demonstrated one form with its Speedfactory for individualised running shoes [9]. Currently, DTMFs exist for three product groups: Knitted fabrics, home textiles, and clothing made from textile surfaces.

The DMTF covers the complete value creating chain comprising all design and production steps from the customer to the ready-made product using these benefits of digitalization and, in most applications or settings, geographical proximity of the production steps. The operations rely on virtual models (digital twins) of the process steps involved, as well as of the materials, the products, and the customers. This digitalization allows effects regarding **flexibility** (deep consumer interaction), **speed** (availability), **efficiency** (less physical waste), **high quality** (no ramp up process), and **personalization** (lot size one). These effects will be picked up in section 3 and assigned there to specific application cases, in order to show their relevance and possible influence, besides the resource savings by using digital twins as long as possible (described in section 3).

In general, the traditional textile production workflow involves manual, labour-intensive steps from creating a design to sewing the final garment [4]. The DTMF approach reduces the long, time-consuming production steps to a minimum, especially in the development step, and ensures a continuous integrated workflow, e.g. with the help of the following value creation steps in a DTMF that uses digital textile printing on textile surfaces as a core process:

- **Body Scanning:** There are already 3D body scanning technologies that allow the generation of digital twins and a virtual try-on in a 3D simulation programme. This can be used for made-to-measure production and can help to determine body data on demand in order to integrate measurement data into existing processes. This enables the body scanner to replace the tape measure with digital measurements on an individual avatar. As a result, the measurement data can be reused and compared in the long term [10].
- **3D/Design**: The basis of a DTMF is the development of creative design in CAD, where garments with real materials (colours/texture) are digitally mapped, e.g. onto individual avatars (virtual models). Thus, in this design step avatars that were previously created with the body scanner can be imported to adapt and grade cuts to individual measurements [11]. The realistic representation makes the interaction of material, cut and body in particular visible (e.g. virtual fit analyses) [12] which can save samples and thus development costs and effort to a considerable extent. With the help of a 3D simulation, the design is then prepared for cutting out [11].
- **Raster Image Processing (RIP)**: After the 3D design, a ‘print and cut’ file is created, which contains a multi-layer file and displays different layers, for example for contours and textures. For this purpose, identifying QR codes and position markers are integrated into the production order for later position recognition. It facilitates the colour-compliant preparation of the design data for the digital printer [11]. This workflow enables more efficient and resource-saving production compared to traditional ones. It is also suitable for on-demand productions, as the workflow can be flexibly adapted.
- **AR/VR**: Virtual interaction options such as Augmented and Virtual Reality (AR/VR) can also be integrated into the DTMF, representing a digital showroom – a virtual place where collections can be viewed virtually. There are already digital and interactive 3D product presentations that can be used at the point of sale [13]. This saves the time-consuming procedure of handing over collections (via photo creation, manual uploading of files, entering in different tools) as it has been practised in previous processes.
- **Digital Printing**: In the next step, the textiles are printed with the individual designs using the core process, the digital printing process [14]. The production files required for this are generated directly from a 3D simulation environment. The colour-accurate preparation of the design data is made possible with the help of the aforementioned RIP programme. The designs are printed on transfer paper, using sublimation equipment. The following thermo-fixing process is accomplished using the calender, thus ensuring a perfect print [11].
- **Cutting Out:** With the help of a camera, the identifying QR codes and position markers make it possible to identify the exact position of each component and the material then can be cut out, entirely automatically [11].
- **Handling**: The cut outs then can be sorted, in a completely automated process, using a robotic arm with a grab claw transporting them, directly and as efficiently as possible, to the assembly department [11].
- **Assembling**: Finally, the individual components are joined together to form a finished product, e.g. by means of sewing or ultrasonic welding machines [11].

3. Business Models for DTMF

A DTMF bears great innovation potential in a broad range of applications and Business Model scenarios, covering B2B settings, where they can improve the prototyping phase of product development up to B2C settings for the innovative production of single-lot, on-site, on-demand individualized products [15,16]. Those benefits of DTMF so far are not widely seen and realised in the TCI due to perceived investment risks and uncertainty with regard to new Business Models. Although there are different generic approaches for Business Model Innovation [17,18] the specific characteristics of the TCI call for a closer look at potential DTMF Business Models.

The economic use of DTMF is a dynamic problem and depends on external conditions and internal design fields and their mutual influences. External conditions include product requirements and dominant market mechanisms, such as trends, competition, demand and delay effects. The internal design fields also include a large number of aspects, such as the setting of the DTMF, the necessary number of machines per value-added stage (scaling), the possible throughput of the DTMF (dimensioning), the operational classification of the DTMF and the associated unit costs and margins to be realised, the resulting throughput times and the competence and personnel requirements.

Besides the economic dimension of sustainability of DTMF – leading to reduced costs and increased margins – there are ecological aspects that refer mainly on '… reduction of waste and transportation needs' [15]. Even digital textile printing as a core process and standalone technology is able to reduce the consumption of ink, waste, energy, and water in comparison to screen printing [15] and offers for all applications potentials to resource saving. The social dimension is met by the local or regional production being close to the customer, focusing together with the data-based transparency on socially responsible innovation and production [15].

Using the DTMF as an integrated concept covering all value creation steps in one place, as described in section 2 (possibly excluding the body scanning), there are at least four application cases that can serve for Business Models, as indicated above: individualization, sampling, reordering and event-driven and locally centred production. Those application cases have been developed in European and German research projects [2,14]. Especially the following ones, labelled as **individualization**, **sampling** and **event-driven and locally centred production** are an outcome of the European Research project 'A Knowledge-based business model for small series fashion products by integrating customized innovative services in big data environment', where they have been derived from different production scenarios [2]. The application case **reordering** has been developed in the wake of German research projects, dealing with the digital transformation in the apparel and clothing industry [14]:

- **Individualization**: Companies can participate in the megatrend of individualization, which is estimated to have a positive impact on turnover next 10-15 years [19], even though there are still few quantitatively reliable statements on turnover development. Nevertheless, cautiously estimated turnover increases of 0.5% to 1% per year are assumed here. Specific customer requirements and wishes can thus be implemented quickly and efficiently in production. This leads to interactive and customer-driven value creation processes. Individual product designs also provide the customer with a special experience. As digitalization increases the range and choice of products and technologies, and consumer demands for uniqueness, quality and service grow. Accordingly, this can again lead to increased competition for individual products and cost pressure [14] and addresses more or less all three sustainability dimensions. Effected in a positive manner are **flexibility** (deep consumer interaction), **speed** (availability), **high quality** (no ramp up process) and **personalization** (lot size one).
- **Sampling**: Working within a digitalised process, using digital twins as long as possible instead of physical samples offers significant cost reductions in new product development. Recent

investigations within the framework of the AiF research project ‘Digital Collection Development’ (IGF project 20892 N) at the Center of Management Research of the German Institutes of Textile and Fiber Research Denkendorf (DITF) in the area of small series in the workwear sector show saving potentials in the new product development and sample production process of 10% to 20% for a product line (e.g. a specific pair of work trousers) [20]. These savings result from the elimination of several physical (fabric) samples and the faster recognition and elimination of undesirable developments by simulating matching repeats, colours etc. Savings can even double when considering new product lines to be developed. Thus, noticeable cost savings of 20% to 40% are possible in product development, especially in sampling, which can then account for up to 1% of total turnover. This in turn simplifies communication and coordination processes between those involved within the process and thus reduces environmental impact by producing fewer physical samples. Digital sampling, however, also means changes from the perspective of the employees, including training, instruction and qualification as well as high investment costs for new systems (e.g. 3D tools) [2]. Business Models here can lead to services offering to re-invent parts or the whole sampling process and effect in particular **efficiency** (less physical waste) and **high quality** (no ramp up process) addressing mainly economic and ecologic dimensions of sustainability.

- **Reordering**: In the case of sold out and highly demanded products a fast and flexible response would allow to satisfy customers’ needs and profit from them. This could include small batches from lot size one up to a big number of products in order to keep the customer’s loyalty and avoid overproduction in times of unstable demand. Hence, the benefits here are twofold: the production could be oriented due to a conservative forecast or demand with less warehousing costs, and the possibility to retain the customer, who will wait for the desired product, which will convert a lost sale into a so-called backlog [21]. Hence, **speed** (availability) and **high quality** (no ramp up process**)** are mainly addressed in a positive way and contribute strongly to economic dimensions of sustainability.
- **Event-driven and locally centred production**: Concerning several events, such as seasonal, sport and cultural ones, there is an opportunity to sell products – often quite simple ones, like T-shirts or scarfs – to customers being fans, supporters, showing to public their opinion etc. This demand is often not predictable and does not allow long delivery times. This asks for a production in the surrounding area, a local production, with a fast and flexible response time. There is in general no time for ramp-up and complicated design and accessories, as the time to market is very close.
  The locally centred production, not only for events, as it is proclaimed in the trend of nearshoring [22], allows to react fast and flexible by using known structures, common rules and conditions [2]. This Business Model requires a powerful infrastructure (using digital twins) and a lot of know-how as well as qualified partners and workers. It could increase the risk of losing touch with new developments due to missing global influences. This means that innovative products must constantly be brought to market and organizational structures must be adapted [2]. In addition, there is the opportunity to choose resources in an environmentally oriented way, to produce under local conditions and thus to comply with socio-ecological standards, and the extent of product counterfeiting can be reduced, which often is a huge problem [23]. This effects positively **flexibility** (deep consumer interaction), **speed** (availability), **efficiency** (less physical waste), **high quality** (no ramp up process), and **personalization** (lot size one). Furthermore, all dimensions of sustainability are addressed to some extent here.

These Business Models can be supported by the following settings of the DTMF [24]:

- **Factory-in-Shop**: A DTMF placed in a retail and/or selling environment focusing on customer or consumer interaction with a fast throughput and production time.

- (**Standalone) Factory:** A DTMF with upscaling capacities allowing to produce a fast and flexible on-demand production, e.g. high-speed printing using multiple printers.
- **Factory-in-Factory:** A DTMF as a workplace in a textile or garment factory for dedicated production jobs (sampling, lot-size one).
- **Technological Centre (Lab):** A DTMF as part of a technological centre or a lab (following the fab lab concept) for design, experimenting, co-creation of products as well as training and education purposes of processes.

In general individualization (1) can be supported by 'Factory-in-Shop' and '(standalone) Factory', sampling (2) by 'Factory-in-Factory' and 'Technological Centre (Lab)', reordering (3) by '(standalone) Factory' and event-driven and locally centred production (4) by 'Factory-in-Shop' and '(standalone) Factory' as shown in Table 1:

Table 1: Settings of DTMF for different Business Models

Business Model	Positive Effects (in addition to resource saving) on	Dimension of sustainability addressed	DTMF Setting	Type of Business Model
Individualization	Flexibility, speed, high quality, personalization	Ecologic, economic, social	Factory-in-Shop, (standalone) Factory	B2C
Sampling	Efficiency, high quality	Ecologic, economic	Factory-in-Factory, Technological Centre (Lab)	B2B
Reordering	Speed, high quality	Economic	(Standalone) Factory	B2C
Event-driven and locally centred production	Flexibility, speed, efficiency, high quality, personalization	Ecologic, economic, social	Factory-in-Shop, (standalone) Factory	B2C

These settings allow in general to use the DTMF as integrated process including all steps presented in section 2 in one place for a fully networked production from body scanning, 3D simulation of the individual garment to digital printing and cutting out to the finished product. In addition, there are other settings (especially sampling) that allow e.g. the production of samples in digital networked forms widely spread over different countries, using only parts of the integrated concept DTMF [24].

Encouraged by the DTMF economic benefits, a trend towards in-house manufacturing for several fashion retailers can already be observed (using the Factory-in-Factory setting), to control their supply chain and to have the benefit of speed to market as well as sustainability [25].

4. Summary

Recent trends and developments in the TCI call for an increased use of digital technologies in order to address changing market and sustainability needs. In this paper we have shown potential Business Models for a fully networked DTMF that uses digital textile printing as a core process step, starting with body scanning, 3D/Design, RIP, AR/VR and completing with cutting out, handling and assembling.

There are four scenarios (or application fields) for Business Models in the areas of B2B and B2C that can use a DTMF in order to cope with the challenges and yield profit. Therefore, the positive effects on the scenarios are indicated and assigned to each of them as well as the dimensions of sustainability. The demand of customers for personalised products with high quality leads to the most noticeable application, the individualization that can be answered by fast local value chains as well as by organizational structures that create new opportunities through end-to-end digitalization. Besides individualization of products (in small lot sizes), there are other promising applications and Business Models for sampling, reordering and event-driven as well as locally centred production, which benefit in the same way from digitally networked end-to-end digital design and production processes. Different settings of DTMF match the presented Business Models.

As an integrated and local production site the DTMF offers a fast (reducing the time to market), flexible and sustainable answer, especially for the fashion and clothing industry. But it has to be adapted according to the Business Models and settings with regard to the specific external conditions (product requirements and dominant market mechanisms, such as trends, competition, demand and delay effects), as well as the internal design fields (like setting of the DTMF, number of machines per value-added stage the dimensioning, personnel requirements) and their mutual influences in order to reach out for the greatest possible economic and sustainable success.

Acknowledgements

We would like to thank the European Union for funding the research project ‘A Knowledge-based business model for small series fashion products by integrating customized innovative services in big data environment’ (Grant Agreement 761122, http://www.fbd-bmodel.eu/) in Horizon 2020 research and innovation programme under grant agreement N. 761122.

In addition, we would like to thank the Project Management Agency Karlsruhe (PTKA) for the support and the Federal Ministry of Education and Research (BMBF) for funding the research programme ‘Retail 4.0 - Virtual Retailer feedback and augmented shopping as a knowledge- intensive, digital service in the development and sales process’ (funding code 02K16C231) and we would like to thank the German Federation of Industrial Research Associations ‘Otto von Guericke’ eV (AiF) for the financial support of the research program ‘Digital collection development - Selective digitization of collection development processes in the clothing industry on the basis of company-specific potential analyzes using parameterizable MFCA models’ (IGF project number: 20892N), which was funded by the Federal Ministry of Economics and Technology (BMWi).

References

[1] Kuhl, J., Krause, D., 2019, Strategies for Customer Satisfaction and Customer Requirement Fulfillment within the Trend of Individualization, Procedia CIRP, Volume 84, 130-135.

[2] Winkler, M., Stellmach, D., Tilebein, M., 2019. Neue Szenarien der Wertschöpfung für Geschäftsmodelle in der Textilwirtschaft, in: Schröder, M. (Ed.), Logistik im Wandel der Zeit – Von der Produktionssteuerung zu vernetzten Supply Chains, Springer Fachmedien, Wiesbaden, pp. 601–626.

[3] Curtis, S.K., Mont, O., 2020. Sharing economy business models for sustainability, Journal of Cleaner Production, Volume 266, 121519, 1-15.

[4] Lushan, S., 2018, Technology disruptions: exploring the changing roles of designers, makers, and users in the fashion industry, International Journal of Fashion Design, Technology and Education, Volume 11, Issue 3, 362-374.

[5] Kiel, D., Müller, J.M., Arnold, C., Voigt, K.-I., 2019. Sustainable Industrial Value Creation: Benefits and Challenges of Industry 4.0, International Journal of Innovation Management, 21(8), 1-34.

[6] Federal Ministry for Economic Affairs and Climate Action, 2022. What is Industrie 4.0?, URL: https://www.plattform-i40.de/IP/Navigation/EN/Industrie40/WhatIsIndustrie40/what-is-industrie40.html, (28.01.2022).

[7] Montes, J., Olleros, F., 2019. Microfactories and the new economies of scale and scope, Journal of Manufacturing Technology Management, 31 (1), 72-90.

[8] Mishima, N. Tanikawa, T. Ashida, K., Maekawa, H., 2002. Design of a Microfactory, in: 7th Design for Manufacturing Conference, Tumer, I. Y. (Ed.), American Society of Mechanical Engineers, New York, pp. 103-110.

[9] Adidas Group, 2019. adidas deploys Speedfactory technology at Asian suppliers by end of 2019, URL: https://www.adidas-group.com/en/media/news-archive/press-releases/2019/adidas-deploys-speedfactory-technology-at-asian-suppliers-by-end-2019/ (14.01.2022).

[10] Shen Y., 2020. 3D Technology and Tailored Clothing, 3rd International Conference on Global Economy, Finance and Humanities Research, European American Chamber of Commerce and Industry, URL: https://www.clausiuspress.com/conferences/LNEMSS/GEFHR%202020/GEFHR2020022.pdf (20.12.2021).

[11] DITF Denkendorf, 2017. MICROFACTORY 4 Fashion, URL: https://www.ditf.de/en/index/more-informations/microfactory.html (20.12.2021).

[12] Lin Y.-L., Wang M.-J., 2014. Digital Human Modelling and Clothing Virtual Try-on Proceedings, Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management, Bali, Indonesia, January 7 – 9, 1914-1918.

[13] Vuframe, 2021. Unleash your products with interactive 3D, URL: https://en.vuframe.com/ (20.12.2021).

[14] Moltenbrey F., Fischer T., 2021. Retail 4.0 – Digital Customer and Retailer Feedback to Garment Development, Journal of Textile Science & Fashion Technology, 8 (2), 1-4.

[15] Tilebein, M., 2019. Small, Smart and Sustainable, Digital textile (2), World Textile Publications Ltd, Leeds, UK, 60-63.

[16] Tilebein, M., 2017. New Business Models, IoTex – Technical Brief, Issue 1, 38-39.

[17] Johnson, M. W., Christensen, C. M., Kagermann, H., 2008. Reinventing Your Business Model, Harvard Business Review, 86 (12), 57-68.

[18] Osterwalder A., Pigneur, Y., 2013. Business model generation: A handbook for visionaries, game changers, and challengers, Wiley&Sons, New York.

[19] Peters R., Goluchowicz, K., Richter S., 2020. Perspectives 2035: A guide to the textile future | Summary, URL: https://textil-mode.de/en/research/zukunftsstrategie-perspektiven-2035/ (14.01.2022).

[20] AiF Datenbank, 2022. Digital Collection Development – Final report of the research project (IGF No. 20892 N), URL: https://www.aif.de/foerderangebote/igf-industrielle-gemeinschaftsforschung/igf-projektdatenbank.html (27.01.2022).

[21] Absi, N., Kedad-Sidhoum, S., Dauzère-Pérès, S., 2011. Uncapacitated lot-sizing problem with production time windows, early productions, backlogs and lost sales, International Journal of Production Research, 49:9, 2551-2566.

[22] Anderson, J., Berg, A., Hedrich, S., Ibanez, P., Janmark, J. and Magnus, K.-H., 2018. Is apparel manufacturing coming home? Nearshoring, automation, and sustainability – establishing a demand focused apparel value chain, McKinsey & Company, URL: https://www.mckinsey.com/~/media/mckinsey/industries/retail/our%20insights/is%20apparel%20manufacturing%20coming%20home/is-apparel-manufacturing-coming-home_vf.pdf (20.12.2021).

[23] Meraviglia, L., 2018. Technology and counterfeiting in the fashion industry: Friends or foes?, Business Horizons, Volume 61, Issue 3, 467-475.

[24] Artschwager, A. Winkler, M., Brunner, L., 2022. Textile Microfactory and Distributed Production – Digitalisation as a Game Changer for Backshifting – A response to changing environment and consumers. Why Textile Microfactories and Digitalisation will change the Future of Textiles, White Paper, URL: www.ditf.de/ (22.03.2022).

[25] McKeegan, D., 2018. Why Textile Micro-Factories Will Change the Future of Fashion, Which PLM, URL: https://www.whichplm.com/why-textile-micro-factories-will-change-the-future-of-fashion/ (20.12.2021).

Biography

Marcus Winkler (*1965) studied economics at the University of Stuttgart and has been working for the Center of Management Research of the German Institutes of Textile and Fiber Research Denkendorf (DITF), since 1995. The research of Dr. Marcus Winkler includes Internet 4.0, Supply Chain Management and Business Models in national and international research and consulting projects.

Franziska Moltenbrey (*1987) has been working for the Center of Management Research of the German Institutes of Textile and Fiber Research Denkendorf (DITF) since 2017. She studied textile management and retail research at Reutlingen University (M.Sc.) and works at the DITF in the fields of digitalization, collection development and microfactories within the texile and retail industry.

Meike Tilebein (*1966) has been the Director of the Institute for Diversity Studies in Engineering at the University of Stuttgart since its establishment in 2009. Prof. Dr. rer. pol. Dipl.-Ing. Meike Tilebein also heads the Center of Management Research of the German Institutes of Textile and Fiber Research Denkendorf (DITF) that supports the textile industry’s digital transformation.

3rd Conference on Production Systems and Logistics

Developing Gaia-X Business Models For Production

Felix Hoffmann[1], Markus Weber[1], Matthias Weigold[1], Joachim Metternich[1]

[1]*Institute of Produciton Management, Technology and Machine Tools(PTW)/ TU Darmstadt, Darmstadt, Germany*

Abstract

The manufacturing industry is in the midst of a digital transformation. As part of the increasing internal and external integration of manufacturing companies, ever more significant volumes of data are being exchanged in order to meet the challenges of a globalized production world. The European initiative Gaia-X aims to establish a federal data infrastructure based on European law to ensure data sovereignty in the resulting digital value creation ecosystems. Under the conditions thus created, it will be possible for manufacturing companies to develop entirely new business models. Within the scope of these business models, the benefit of data sharing in the sense of added value will come into focus.

The following paper presents opportunities for the development of disruptive digital business models for manufacturing companies in the context of Gaia-X. The paper focuses on how data sharing can be used to create value. Furthermore, it highlights how the transition from technological use case to monetizable value creation can be made with data-based, digital business models in the context of Gaia-X. Finally, the state of work in business model development in the Gaia-X project EuProGigant is presented for discussion and exemplified by two use cases.

Keywords

Digitization; Business Models; Gaia-X; Production Technology

1. Introduction

The industrial sector is currently in the midst of a fundamental digital transformation. In the last 12 years, the amount of data generated worldwide almost increased by fifty, and progressive growth is expected in the coming years as well [1]. Manufacturing companies support this increase by constantly driving forward the digitization of their products and processes. Due to the advancing use of sensors and increasing connectivity of machines and systems, information availability continues to rise [2]. In this context, digital, data-based business models represent an essential foundation for generating benefits from the data acquired. A platform-based exchange of data across locations and company boundaries becomes increasingly important as a key benefit driver [3]. However, the spread of such platform-based business models is very limited. Many potential players are not willing to participate out of fear of losing their data sovereignty [4]. The European initiative Gaia-X, launched in 2019, addresses this challenge by establishing a federated data infrastructure to ensure data sovereignty based on the European legal situation. Gaia-X's decentralized approach aims to aggregate the heterogeneous infrastructures of different actors into a homogeneous system. Those systems are named ecosystems or data space and are characterized by technology, business and legal [5]. In this context, the idea of open source is of high priority. Especially smaller companies can also benefit from the development. Trust is established through transparency of code, contract and verifiable identities and credentials. Furthermore, various instances are networked via open interfaces and standards to optimize the

DOI: https://doi.org/10.15488/12130

linkage of data sources and sinks. This intends to increase the sovereignty of customers of platform-based business models and the scalability, interlinkage and competitive position of their providers [6].

Research on platform-based business models in industrial production and the accompanying empowerment of companies is still in its infancy [3]. The same applies to the development efforts on building data infrastructure in the Gaia-X context. Among other things, parties involved are intensively working on designing the underlying architecture in numerous working groups. Accordingly, the overall requirements for corresponding data-based business models continue to change. The following remarks reveal future possibilities of platform-based business models within Gaia-X. First of all, this paper addresses the concept of platform-based business models. Subsequently, it presents a possible procedure for the structured development of decentralized, multi-platform-based business models in the context of Gaia-X. Thus, the current work on developing use cases and business models in the EuProGigant project is addressed and exemplified by two use cases. Due to the early stage of the project and the limited scope of the paper, the application is focused on the initial area of solution development.

2. State of the Art

A business model captures value and generates profitable outcomes through applying a particular technology. A business model is a connecting link between technology and its economic value characterized by the three complementary dimensions of value generation, value proposition and revenue structure [7]. The value proposition dimension depicts the benefits a company offers its customers with a particular product or service. The value generation dimension captures central processes and competencies required to implement the business model and fulfil the value proposition. Finally, the revenue structure dimension describes the composition of cost and revenue mechanisms and the resulting value generated from the business [8].

Data-driven, digital business models represent a specific form and have a customer-oriented, service-driven value generation based on data and a full digitalized implementation [9]. Concerning value generation, a data value chain significantly shapes the interactions in such a business model's ecosystem [10]. The data thus utilized can be obtained from various internal and external data sources [11]. In the manufacturing field, data often originates from using products such as machine tools. This is not least due to the ongoing transition from physical products to product-service systems and software-as-a-service models, as the significance of dematerialized value increases continuously [12]. There is also an adjustment in the profile of the players involved in a data-driven business model — the three essential roles of data user, data supplier and data enabler—the three essential roles of data user, data supplier and data enabler [13]. The data user utilizes the data resources available to him in order to create and realize value. The value creation can focus on internal and external value creation (optimization of internal process vs sale of products). The data supplier or data enabler supports the data user in his activities. A data provider ensures a supply of context-specific, relevant data. In contrast, a data enabler provides supporting data services or data infrastructure solutions [13,14]. The interaction of these players is not characterized by one-off or sporadic interactions but by reoccurring and regular ones. Accordingly, there is also a change in the revenue structure to reflect this transformation of service exchange. Thus, the trend is toward repetitive transactions in data-based service bundles. This trend includes subscriptions, key figure-oriented billing (e.g., payment per component produced) or profit-sharing (e.g., participation in savings achieved through using a product). Likewise, compensation models are conceivable in which payment is made through the provision of data [15].

Platform-based business models pick up on this aspect of a transformation in exchanging goods and services and drive it further. Their goal is to reach a more significant number of different participants and facilitate interactions between them [16]. In the business-to-consumer sector, such digital platforms are already widespread. A fundamental distinction can be made between three types of platforms: aggregation, social and mobilization platforms. Aggregation platforms merge a wide range of relevant resources. They help a user connect to the resources he needs, making them highly transactional and task-oriented. The most common examples of this form are broker platforms like eBay and Amazon. Aggregation platforms often operate according to the hub-and-spoke principle, in which a platform owner mediates all transactions. Social platforms aggregate users and support engagement among those with common interests. The most common examples of this form are social media platforms like Facebook or Twitter. Social platforms mainly foster networks without the involvement of an organiser or owner.

Lastly, mobilization platforms get users to collaborate to achieve common goals. Long-term relationships are targeted instead of completing short-term transactions or tasks. Mobilization platforms connect users in extended business processes, such as delivery networks or sales operations. Well-known examples of this are the global supply chain platform Li & Fung or Linux and Apache software platforms [17]. In the context of production, aggregation and mobilization platforms are in focus. In terms of data processing, these two concepts enable capturing financial value from data assets. The data provider and the platform provider can achieve a corresponding monetization. Thus, such platforms position themselves as a central interface between data user, data supplier and data enabler within a cross-process value network [4]. Although the spread of platform-based B2B business models in production is still in its infancy, the first corresponding offerings are already on the market [4]. However, these are essentially proprietary applications from machine manufacturers for company- or lifecycle-phase-specific applications. This contrasts with the openness and trustworthiness of digital platforms as a decisive success factor, as the Gaia-X initiative aims [18].

3. Methodology

The following section addresses how the transition from technological use case to monetizable value creation can be performed within Gaia-X. To this end, the approach to business model development pursued in the EuProGigant project is depicted. The project is a German-Austrian cooperation, which was selected by the Gaia-X initiative as a lighthouse project in the production environment. The presented approach emerges from process models and methods of business model innovation and data science (see Figure 1).

Successful implementation of data-based business models for production requires a systematic and structured process [19]. Concerning the underlying data-based applications, numerous process models exist in the literature. Most of them originate from the field of data mining [20]. Well-known approaches in this field include the Cross-Industry Standard Process for Data Mining (CRISP-DM), the Sample, Explore, Modify, Model, Assess (SEMMA) and the Knowledge Discovery in Databases (KDD) [21]. A deeper analysis of the models in terms of their suitability for the manufacturing industry reveals numerous shortcomings. These prevent a practical and holistic application in such a domain. Among the main criticisms are a lacking possibility of problem selection and a missing consideration of specific requirements from production environments [22].

In order to address these shortcomings, Biegel et al. [22,19] introduced their own Artificial Intelligence Management Model for the Manufacturing Industry (AIMM). Although the model has its bases on artificial intelligence, the approach can also be adapted to the area of platform-based business models. This work then further utilizes the AIMM as a general framework for business model development. In the course of expert workshops in EuProGigant, the model was adapted in broad areas to the already known framework conditions of Gaia-X. This includes, among other things, necessary criteria and building blocks that enable the implementation of business models with Gaia-X.

The process model is funneled and starts with potential problems, subsequently transformed into an application (see Figure 1). The approach has three phases: problem selection, solution design and solution development. In the initial phase of problem selection, the project team first identifies and evaluates relevant problems from the production environment. These are then compared in terms of their complexity as well as relevance and their Gaia-X fit is checked. Promising approaches are selected for further work in the solution design phase. In this phase, the approaches are developed into business models from a holistic perspective. Further, they are evaluated in terms of their technical, organizational and economic feasibility. In the final solution development phase, the elaborated concepts are finally realized, tested and implemented in a development project. A significant difference between the process phases results from the availability of relevant information and the present degree of uncertainty. At the beginning of the process, there is only a low level of information and, at the same time, a high degree of uncertainty. This relationship is reversed as the process progresses [23]. The approach is further designed to fail quickly in the case of an unpromising endeavor. This considers that, particularly at the beginning of an application development process, the efforts incurred are still low. At the same time, a strong influence can be exerted on the future cost-benefit ratio in later phases of development and utilization [24]. Therefore, the process enforces to evaluate if a business case is technically, organizationally, financially and legally – e.g., in terms of data sovereignty – feasible. If an approach drops out, the process can be revisited with a different problem. Otherwise, the solution design can be adjusted accordingly. In this way, the waste of entrepreneurial resources is prevented at an early stage [22].

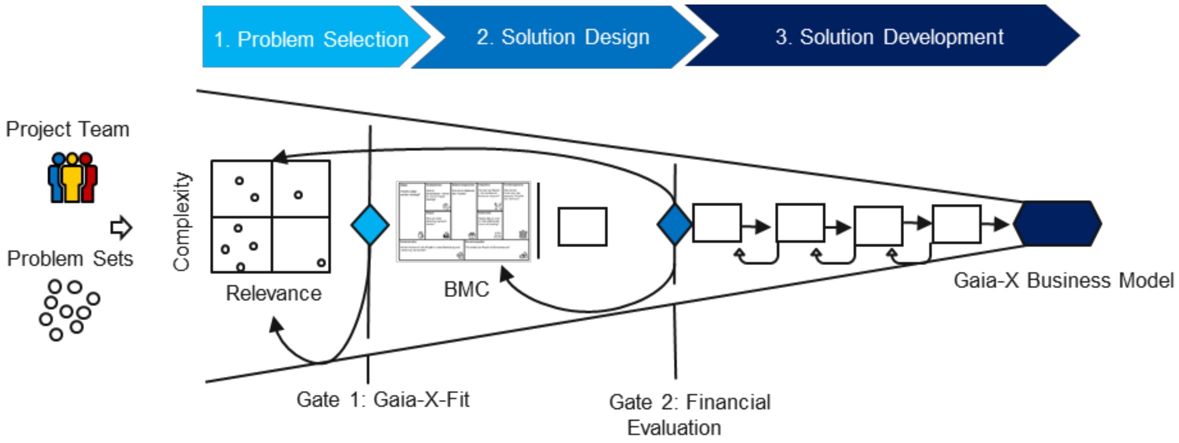


Figure 1: EuProGigant business model development process in accordance with [22,25]

In addition to the evaluation mentioned above within dropout gates, the approach also integrates tools from business model innovation. These are applied in particular in the solution design phase. One of the tools used in this phase of the process model is the Business Model Canvas (BMC) by Osterwalder and Pigneur. The BMC is a framework for visualizing and structuring business models. It is used to generate initial business ideas and creates a holistic overview of business model components. Based on the already presented areas of a business model, the BMC divides them into a total of nine segments, namely: key partners, key activities, key resources, value proposition, customer relation, channels, customer segments, cost structure and revenue streams [25]. The advantage of the BMC is the ability to present a business model in a holistic and clear way and thus to identify possible dependencies. In addition, a uniform understanding of the significance of individual components of the business model can be generated in a project team [26]. One drawback of the model for application to data-based business models is its high degree of generality. Metelskaia et al. [27] address this shortcoming in their extension of the BMC. Based on a comparison of existing approaches to combining business models and data analysis, they specify possible contents of the canvas elements. For example, the key partners are supplemented by IT and data science companies and the revenue streams include novel approaches like Pay-per-X. These specifications make it easier for inexperienced users to create their own approaches with the help of the BMC.

4. Application

In this section, the current work on developing use cases and business models in the EuProGigant project is presented and exemplified in two use cases. Compared to an application in a real industrial environment, there is a significant difference when applied to a research project: Whereas in industry one often must choose between working on different problems arising from one's own company or from customer requirements, the problem in a research project is usually already defined in advance. For this reason, it was decided not to apply the methods from the Problem Selection phase. Furthermore, due to the early stage of the project and the limited scope of the paper, the application is focused on the initial area of solution development. For this purpose, the use case is first described, and then the BMC is applied. The two use cases shown are the ideal component matching and the validation platform. The results presented were developed within interdisciplinary workshops with the project participants. In both cases, domain experts, data scientists, as well as software and electronics developers were among the participants.

4.1 Ideal Component Matching

The assembly of modules (e.g., a shaft-hub connection) combines individual parts from various sources. Typically, some of these components are manufactured in-house by machining companies, and the rest of the parts are sourced from different suppliers. Due to stochastic variations in each company's manufacturing environment, the actual geometries of the components generally deviate slightly from the specifications. Limits are set for combination tolerances of the assembly and allowable deviations of individual parts. Specially manufactured components compensate deviations of a sum tolerance. The solution involves the use of sensory tools and workpiece clamping devices. The data is processed using artificial intelligence methods. In this way, the identification of statistical correlation between component dimensions and processes is enabled. This allows manufacturers to improve the quality of their assemblies and produce targeted matching components as needed.

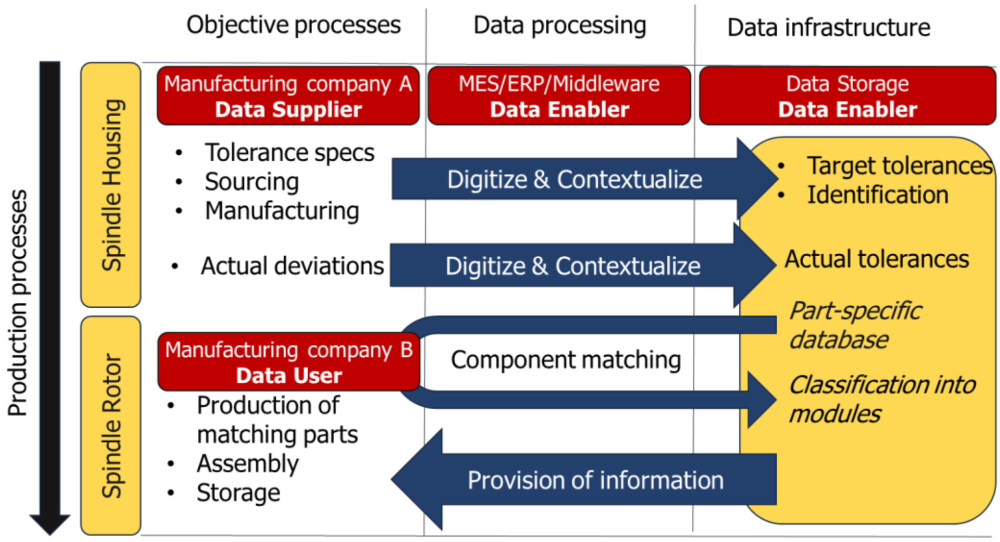
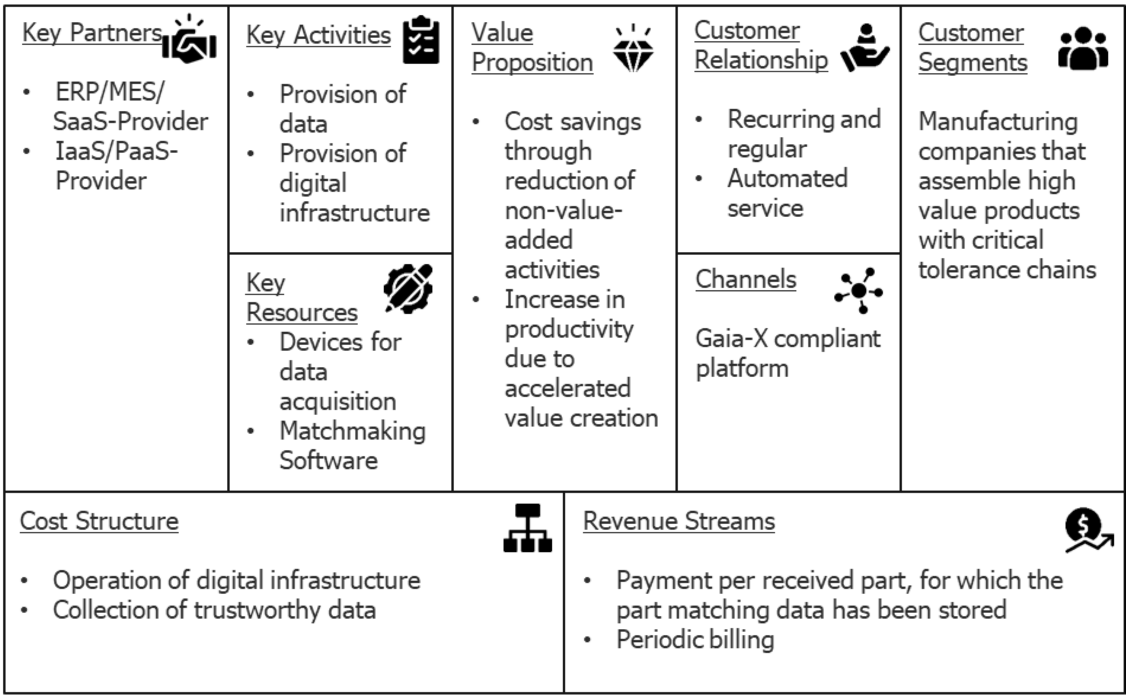


Figure 2: Concept of ideal component matching for a machine tool spindle

In the EuProGigant project, the novel concept is being tested on a machine tool spindle. Two project partners manufacture the two relevant components in the spindle housing and the spindle rotor at different locations. One reason why the machine tool spindle is suitable for the concept is that it is a higher-value component that accounts for a relevant proportion of the total cost of the end product. In addition, the spindle is essential for the manufacturing accuracy and thus for the quality of the components manufactured on a machine tool [28]. Therefore, there are high requirements for the manufacturing accuracies of the housing and the rotor. The same applies to the fitment accuracy and the concentricity of the resulting assembly. The concept of ideal component matching in EuProGigant is shown in Figure 2. The concept is only possible by the close interaction of data supplier, data enabler and data user. In this process, manufacturer A produces the spindle housing according to the tolerance specifications. The captured data - including actual deviations - is then digitized and contextualized via the middleware and stored in the data storage. The stored data are uniquely assigned to each produced component. With the available data, manufacturer B can identify the ideal counterpart to the spindle rotors it manufactures. In this process, the component data for component matching is merged via the middleware and ideal pairs of spindle and rotor are identified. The result in the form of a classification into modules is finally stored in the data storage. Manufacturer B can then plan its component assembly based on this information. Furthermore, manufacturer B can produce a matching spindle rotor based on this data if no corresponding counterpart is available.



Key Partners
- ERP/MES/ SaaS-Provider
- IaaS/PaaS-Provider

Key Activities
- Provision of data
- Provision of digital infrastructure

Key Resources
- Devices for data acquisition
- Matchmaking Software

Value Proposition
- Cost savings through reduction of non-value-added activities
- Increase in productivity due to accelerated value creation

Customer Relationship
- Recurring and regular
- Automated service

Channels

Gaia-X compliant platform

Customer Segments

Manufacturing companies that assemble high value products with critical tolerance chains

Cost Structure
- Operation of digital infrastructure
- Collection of trustworthy data

Revenue Streams
- Payment per received part, for which the part matching data has been stored
- Periodic billing

Figure 3: BMC applied on the ideal component matching

Figure 3 shows the application of the BMC. A central value proposition of the ideal component matching is a significant reduction in non-value-adding tasks. Another value proposition is a higher resource efficiency due to fewer rejected parts. This value proposition is made possible through a trustworthy data transfer within Gaia-X. The concept eliminates the need for a direct sequence of final goods inspection at the supplier and incoming goods inspection at the customer. Instead, the customer receives trustworthy component information directly from the supplier's final inspection. Furthermore, it enables creating time flexibility potentials in cross-company value chains. Thus, the sustainable value contribution for the stakeholders of the use case ideal component matching lies primarily in an increased speed of value creation. Through a resulting reduction in assembly time, a possible productivity increase of 10% can be achieved in case of the machine tool spindle. The data provider - in other words, the component supplier - and the data enabler - in other words, the infrastructure provider - can be remunerated for this added benefit to the data user within a revenue model. The pricing can thereby be aligned with the expected cost savings per assembled component. In the use case, the payment is made per purchased component for which the matching data was provided during the handover. In this case, billing can take place at regular intervals. This takes the high number of individual contacts and thus transactions into account. Data providers and enablers thus can cover their costs for operating the digital infrastructure and collecting trusted data. Accordingly, they can obtain a profit opportunity as an incentive to participate in the business model.

4.2 Validation Platform

Predictive maintenance in production promises to reduce maintenance costs and the number of unplanned machine downtimes. On the one hand, this can improve economic efficiency. On the other hand, it can increase the availability of machines and systems [29]. Many companies have already recognized the potential of this technology, but they often fail to implement it practically [30]. Predictive maintenance is based on mathematical models, which often originate from machine learning. These models use as input sensor data from machine components both for its training and operation. Especially models that are supposed to predict the remaining lifetime of components depend on a broad basis of historical data for a reliable output [31]. However, especially in the case of components that bear a high proportion of the cost of a machine – such as a machine tool spindle – it can be assumed that long-term recording of data on several, comparable machines is necessary to provide data records of degradation and wear events in sufficient quantity [32]. In particular, small and medium-sized enterprises have problems with the provision of corresponding data sets. One of the reasons for this is that they often only have access to historical data sets that are not very comprehensive or of insufficient quality [33]. In addition, they often have heterogeneous machine fleets that make collecting data on similar machines and their components even more difficult. The use of a validation platform enables monitoring machines and assemblies for companies without an extensive database. Collaborative and predictive maintenance of machines and their components can thus be enabled due to different companies' shared use of data.

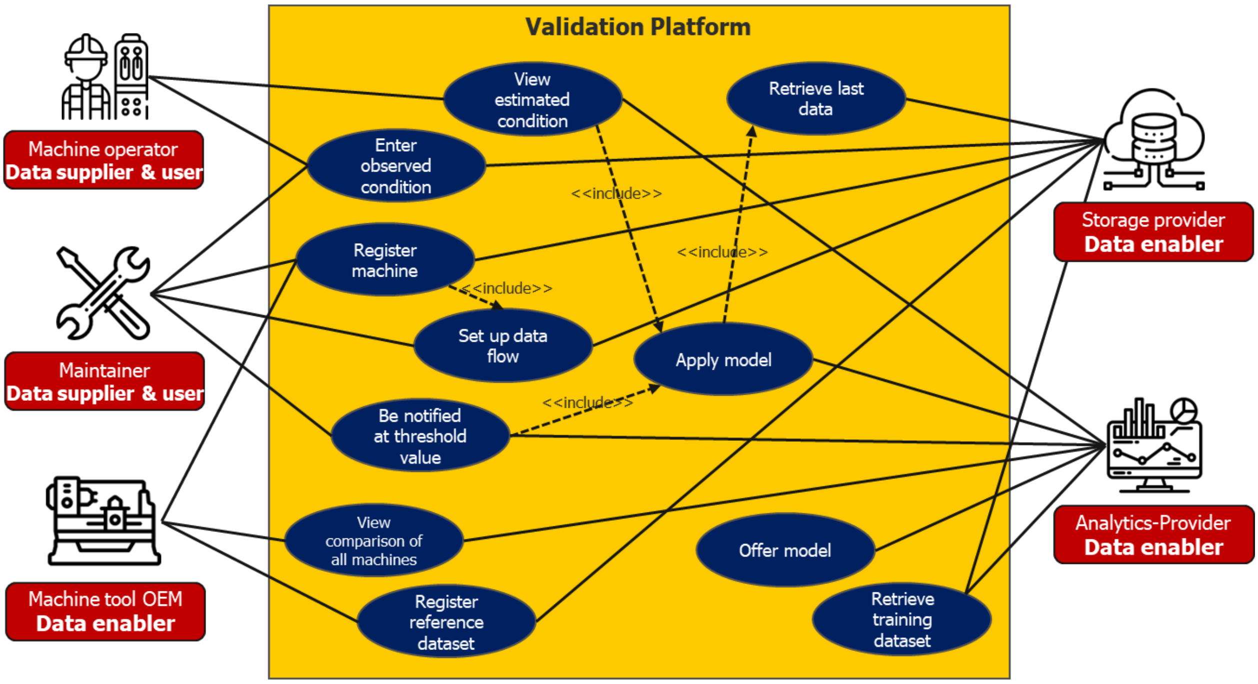
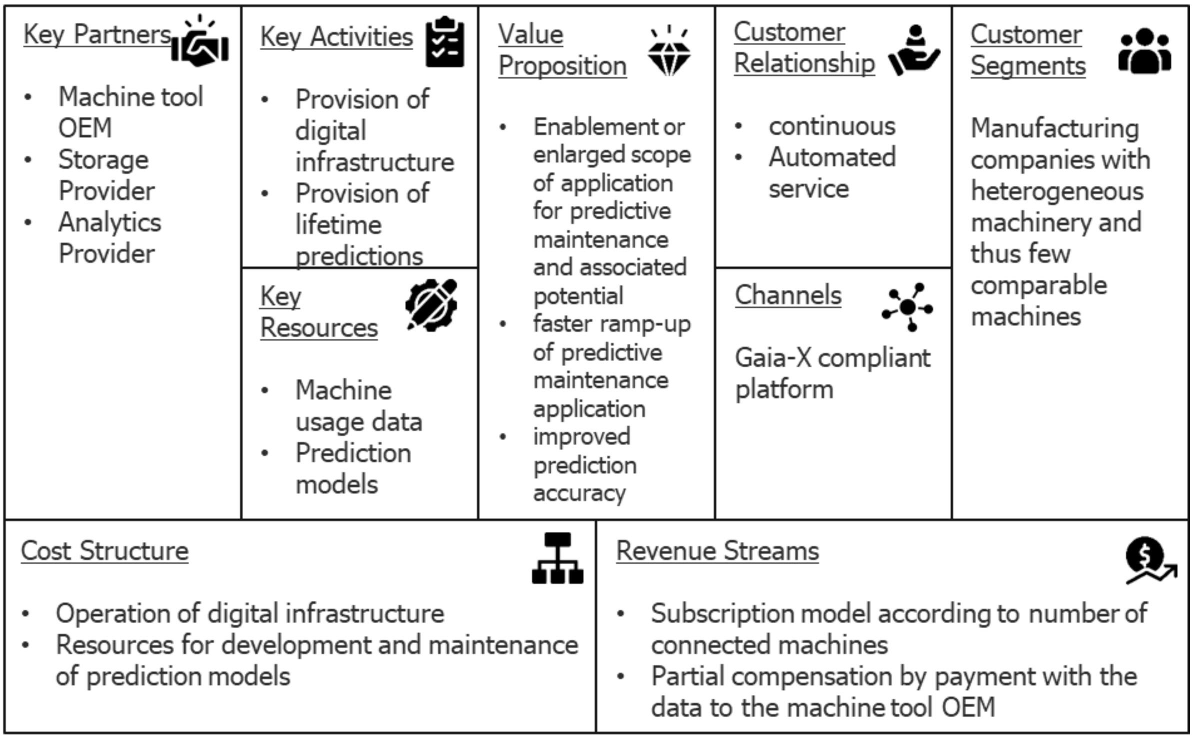


Figure 4: Concept of validation platform for machine tools

The EuProGigant project tests the concept of a validation platform on several similar machine tools. These are located at various sites of different production companies. The concept of the validation platform in EuProGigant can be seen in Figure 4. As in the case of the ideal component matching example, it can be seen here that the approach is only made possible by the close interaction of data suppliers, data enablers and data users. Here, the machine operator and the maintenance engineer simultaneously act as data suppliers and data users. During the operation and maintenance of the machine, both actors generate condition-relevant data, which is stored by the storage provider. The analytics provider can in turn use this data to train and operate its provided condition monitoring model. Thereby the machine tool OEM determines by registration of the machine, which reference data set of similar machine can be used. Based on the results of the condition monitoring model, the machine operator receives an assessment of the machine condition via the platform. Furthermore, the maintainer is informed as soon as the remaining service life of a component falls below a threshold value. Both actors then return feedback regarding the observed condition to the platform. This feedback can in turn be used to improve the condition monitoring model. The machine tool OEM, the storage provider and the analytics provider thus assume the roles of data enablers.

Figure 5 shows the application of the BMC. The validation platform has several complementary value propositions. On the one hand, it enables companies with a heterogeneous machine park to apply predictive maintenance for a more significant part of their machines. Thus, it leverages the potential already presented. Furthermore, it is possible to build up an adequate database more quickly and thus reduce the start-up phases of corresponding solutions. Finally, the prediction accuracy of the models can be improved by a broader data basis with actual process data from machine operation.



Key Partners

- Machine tool OEM
- Storage Provider
- Analytics Provider

Key Activities

- Provision of digital infrastructure
- Provision of lifetime predictions

Key Resources

- Machine usage data
- Prediction models

Value Proposition

- Enablement or enlarged scope of application for predictive maintenance and associated potential
- faster ramp-up of predictive maintenance application
- improved prediction accuracy

Customer Relationship

- continuous
- Automated service

Channels

Gaia-X compliant platform

Customer Segments

Manufacturing companies with heterogeneous machinery and thus few comparable machines

Cost Structure

- Operation of digital infrastructure
- Resources for development and maintenance of prediction models

Revenue Streams

- Subscription model according to number of connected machines
- Partial compensation by payment with the data to the machine tool OEM

Figure 5: BMC applied on the validation platform

A Gaia-X-compliant platform enables trustworthy data transfer and merges data streams from different companies. This ensures that only authorized players can access the data and that there is no leakage of intellectual property over the data from machine usage. The data enablers - i.e., the machine tool OEM, the storage provider and the analytics provider - can generate new cash flows via an appropriate revenue model in return for the added value of the data user. Due to the continuous provision of services, a subscription model is recommended. In the context of the use case, it is intended that payment will be made per connected machine or component. Tiered pricing is also considered a possible model if several machines are connected. This pricing can be based on the expected cost savings due to an enabled or improved predictive maintenance use.

Furthermore, the machine tool OEM can use the data to optimize its own products and product-service offerings. In return, a part of the payments could be compensated by this benefit. Through the revenue streams thus realized, the data enablers have the opportunity to cover their costs of operating the digital infrastructure and developing and maintaining the predictive models. Ultimately they receive a profit opportunity as an incentive to participate in the business model.

5. Conclusion and Future Research

This paper presents a possible methodological approach for developing digital, platform-based business models in the context of Gaia-X. First, fundamental properties of data-driven, platform-based business models were discussed and then a process model was derived. The presented approach and the tools contained therein are practically applied in the context of the Austrian-German lead project for Gaia-X in the manufacturing industry called EuProGigant. Two of the business models considered in the project were finally presented and captured in a BMC, which was utilized in the solution design phase of the project. A key insight from the presentation of the two use cases is that the utility value of a common data infrastructure does not only lie in the direct selling and buying of data and services. It is instead in the saving of value-destroying sections of process chains. These, in turn, open up time-transparent flexibility potential and thus strengthen resilience in the network.

In the considerations made in the context of this paper, it should be noted that the contents presented provide an initial outlook on the future possibilities of platform-based business models within Gaia-X. The Gaia-X initiative and the lighthouse project EuProGigant, are still in their infancy and are currently characterized by high development dynamics. Once the Gaia-X community has created a robust framework, the business models' technical details can be further refined. Thus, the presented process model shall be regarded as a working status. It will be continuously adapted by the progress of the project and optimized and extended with regard to the knowledge gained. Furthermore, the approaches to business model development must be further tested, and their technical feasibility must be confirmed. In the course of this, the evaluation methods outlined can also be used to assess the economic viability of the business models. Adjustments can be made as part of an iterative improvement process if necessary. Lastly, only one section of the process model, namely solution development with the BMC, was considered in the context of the paper. The aim of further work and publications in the project should be to apply and evaluate the tools of the other phases in practice as well.

6. Acknowledgements

The authors thank the German Project Management Agency of the German Aerospace Center (DLR-PT) and the Austrian Research Promotion Agency (FFG) for financial support and project monitoring. The research project EuProGigant is funded in the course of the bilateral funding call "Smart and Sovereign Use of Data for Production" (Big Data in Production), within the Austrian program "FTI Offensive Big Data in Production" of the Research Promotion Agency (FFG) on behalf of the Federal Ministry for Climate Protection, Environment, Energy, Mobility, Innovation and Technology (BMK) and the German program "Development of Digital Technologies" of the Federal Ministry of Economics and Climate Protection (BMWK).

7. Literature

[1] See, A. von, 2021. Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025. https://www.statista.com/statistics/871513/worldwide-data-created/. Accessed 1 February 2022.

[2] Schmitz, C., Tschiesner, A., Jansen, C., Hallerstede, S., Garms, F., 2019. Industry 4.0: Capturing value at scale in discrete manufacturing, Frankfurt am Main. https://www.mckinsey.com/~/media/mckinsey/industries/advanced%20electronics/our%20insights/capturing%20value%20at%20scale%20in%20discrete%20manufacturing%20with%20industry%204%200/industry-4-0-capturing-value-at-scale-in-discrete-manufacturing-vf.pdf. Accessed 1 February 2022.

[3] Plattform Industrie 4.0, 2021. Fortschrittsbericht 2021: Industrie 4.0 gestalten. Wenn Vision Realität wird., Berlin. https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/2021-fortschrittsbericht.pdf?__blob=publicationFile&v=16. Accessed 1 February 2022.

[4] Plattform Industrie 4.0, 2020. Datenmarktplätze: Datenmarktplätze in Produktionsnetzwerken, Berlin. https://www.plattform-i40.de/IP/Redaktion/DE/Downloads/Publikation/datenmarktplaetze-in-produktionsnetzwerken.pdf?__blob=publicationFile&v=7. Accessed 1 February 2022.

[5] Halevy, A., Franklin, M., Maier, D., 2006. Principles of dataspace systems, in: Proceedings of the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems - PODS '06. the twenty-fifth ACM SIGMOD-SIGACT-SIGART symposium, Chicago, IL, USA. 26.06.2006 - 28.06.2006. ACM Press, New York, New York, USA, pp. 1–9.

[6] Gaia-X, 2020. Gaia-X: Das europäische Projekt startet in die nächste Phase, Berlin. https://www.bmwi.de/Redaktion/DE/Publikationen/Digitale-Welt/gaia-x-das-europaeische-projekt-startet-in-die-naechste-phase.pdf?__blob=publicationFile&v=18. Accessed 1 February 2022.

[7] Gassmann, O., Frankenberger, K., Choudury, M., 2014. The business model navigator: 55 models that will revolutionise your business. Pearson, München, 387 pp.

[8] Chesbrough, H., 2017. Business model innovation: It's not just about technology anymore. Strategy & Leadership 35 (6), 12–17.

[9] Sharp, M., Sexton, T., Brundage, M.P., 2017. Toward Semi-autonomous Information, in: Lödding, H., Riedel, R., Thoben, K.-D., Cieminski, G. von, Kiritsis, D. (Eds.), Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing, vol. 513. Springer International Publishing, Cham, pp. 425–432.

[10] Otto, B., Jürjens, J., Schon, J., Auer, S., Menz, N., Wenzel, S., Cirullies, J., 2016. Industrial Data Space: Digitale Souveränität über Daten, München. https://www.fraunhofer.de/content/dam/zv/de/Forschungsfelder/industrial-data-space/Industrial-Data-Space_whitepaper.pdf. Accessed 2 February 2022.

[11] Hartmann, P.M., Zaki, M., Feldmann, N., Neely, A., 2016. Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. IJOPM 36 (10), 1382–1406.

[12] Khan, M.A., Wuest, T., 2019. Upgradable Product-Service Systems: Implications for Business Model Components. Procedia CIRP 80, 768–773.

[13] Schroeder, R., 2016. Big data business models: Challenges and opportunities. Cogent Social Sciences 2 (1), 1166924.

[14] Wiener, M., Saunders, C., Marabelli, M., 2020. Big-data business models: A critical literature review and multiperspective research framework. Journal of Information Technology 35 (1), 66–91.

[15] Schüritz, R., Seebacher, S., Dorner, R., 2017. Capturing Value from Data: Revenue Models for Data-Driven Services, in: Proceedings of the 50th Hawaii International Conference on System Sciences (2017). Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences.

[16] Täuscher, K., Laudien, S. (Eds.), 2017. Uncovering the Nature of Platform-based Business Models: An Empirical Taxonomy.

[17] Hagel, J., 2015. The power of platforms, London. https://www2.deloitte.com/content/dam/Deloitte/za/Documents/strategy/za_The_power_of_platforms.pdf. Accessed 2 February 2022.

[18] Gerrikagoitia, J.K., Unamuno, G., Urkia, E., Serna, A., 2019. Digital Manufacturing Platforms in the Industry 4.0 from Private and Public Perspectives. Applied Sciences 9 (14), 2934.

[19] Metternich, J., Biegel, T., Bretones Cassoli, B., Hoffmann, F., Jourdan, N., Rosemeyer, J., Stanula, P., Ziegenbein, A., 2021. Künstliche Intelligenz zur Umsetzung von Industrie 4.0 im Mittelstand: Expertise des Forschungsbeirats der Plattform Industrie 4.0, München. https://www.plattform-

i40.de/IP/Redaktion/DE/Downloads/Publikation/Expertise-Forschungsbeirat_KI-fuer-Industrie40.pdf?__blob=publicationFile&v=3. Accessed 6 December 2021.

[20] Hand, D.J., Adams, N.M., 2014. Data Mining, in: Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L. (Eds.), Wiley StatsRef: Statistics Reference Online, vol. 5. Wiley, pp. 1–7.

[21] Azeved, A., Rojao, I., Santos, M. (Eds.), 2008. KDD, SEMMA and CRISP-DM: A parallel overview.

[22] Biegel, T., Bretones Cassoli, B., Hoffmann, F., Jourdan, N., Metternich, J., 2021. An AI Management Model for the Manufacturing Industry - AIMM.

[23] Samset, K., Volden, G.H., 2016. Front-end definition of projects: Ten paradoxes and some reflections regarding project management and project governance. International Journal of Project Management 34 (2), 297–313.

[24] Ehrlenspiel, K., Kiewert, A., Lindemann, U., 1998. Kostenverantwortung der Produktentwickler, in: Ehrlenspiel, K., Kiewert, A., Lindemann, U. (Eds.), Kostengünstig Entwickeln und Konstruieren. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 5–16.

[25] Osterwalder, A., Pigneur, Y., 2013. Business model generation: A handbook for visionaries, game changers, and challengers. Wiley&Sons, New York, 278 pp.

[26] Frankenberger, K., Weiblen, T., Csik, M., Gassmann, O., 2013. The 4I-framework of business model innovation: a structured view on process phases and challenges. IJPD 18 (3/4), 249.

[27] Metelskaia, I., Ignatyeva, O., Denef, S., Samsonowa, T., 2018. A business model template for AI solutions, in: Proceedings of the International Conference on Intelligent Science and Technology - ICIST '18. the International Conference, London, United Kingdom. 6/30/2018 - 7/2/2018. ACM Press, New York, New York, USA, pp. 35–41.

[28] Stanula, P., Kohn, O., Lang, E., Metternich, J., Weigold, M., Buchwald, A., 2021. Economic assessment of stress-based payment models. Procedia CIRP 103, 182–187.

[29] Hoffmann, F., Brockhaus, B., Metternich, J., 2020. Predictive Maintenance für Schutzabdeckungen: Vom Geschäftsmodell zur Anwendung. WT Werkstattstechnik 110 (07-08), 496–500.

[30] Lundborg, M., Märkel, C., 2019. Künstliche Intelligenz im Mittelstand: Relevanz, Anwendungen, Transfer. WIK GmbH, Bad Honnef.

[31] Liao, L., Köttig, F., 2016. A hybrid framework combining data-driven and model-based methods for system remaining useful life prediction. Applied Soft Computing 44 (1), 191–199.

[32] Bossler, L.F., Rogalski, T., Stanula, P., Lang, E., Kohn, O., Metternich, J., Weigold, M., Krönung, J., Buchwald, A., 2021. Pay-per-Stress – Belastungsorientierte Leasingmodelle im Maschinenbau. Wirtsch Inform Manag 13 (6), 466–475.

[33] Heimes, H., Kampker, A., Bührer, U., Krotil, S., 2019. Potenziale und Hürden von Data Analytics in der Serienfertigung. I40M 2019 (1), 57–60.

8. Biography

Felix Hoffmann, M. Sc. is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2019. His research interests are ML-driven Business Models in manufacturing and Predictive Maintenance.

Markus Weber, M. Sc. is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2017. Mr. Weber is head of research at the Manufacturing Technology department of PTW and project coordinator of EuProGigant.

Prof. Dr.-Ing. Matthias Weigold is the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology with responsibility of the manufacturing and energy efficiency research groups since 2019. Prof. Weigold contributes many years of industrial experience from management positions at Heidelberger Druckmaschinen AG and SAP SE.

Prof. Dr.-Ing. Joachim Metternich is the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Darmstadt University of Technology since 2012. In addition, Prof. Metternich is the spokesman for the Center of Excellence for Medium-Sized Businesses in Darmstadt and president of the International Association of Learning Factories.

3rd Conference on Production Systems and Logistics

Digital Twin Fidelity Requirements Model For Manufacturing

Christian Kober[1*], Vincent Adomat[1*], Maryam Ahanpanjeh[1*],
Marc Fette[1], Jens P. Wulfsberg[1]
[1]*Institute of Production Engineering / Helmut Schmidt University, Hamburg, Germany*
[*]*Authors contributed equally*

Abstract

The Digital Twin (DT), including its sub-categories Digital Model (DM) and Digital Shadow (DS), is a promising concept in the context of Smart Manufacturing and Industry 4.0. With ongoing maturation of its fundamental technologies like Simulation, Internet of Things (IoT), Cyber-Physical Systems (CPS), Artificial Intelligence (AI) and Big Data, DT has experienced a substantial increase in scholarly publications and industrial applications. According to academia, DT is considered as an ultra-realistic, high-fidelity virtual model of a physical entity, mirroring all of its properties most accurately. Furthermore, the DT is capable of altering this physical entity based on virtual modifications. Fidelity thereby refers to the number of parameters, their accuracy and level of abstraction. In practice, it is questionable whether the highest fidelity is required to achieve desired benefits. A literary analysis of 77 recent DT application articles reveals that there is currently no structured method supporting scholars and practitioners by elaborating appropriate fidelity levels. Hence, this article proposes the Digital Twin Fidelity Requirements Model (DT-FRM) as a possible solution. It has been developed by using concepts from Design Science Research methodology. Based on an initial problem definition, DT-FRM guides through problem breakdown, identifying problem centric dependent target variables (1), deriving (2) and prioritizing underlying independent variables (3), and defining the required fidelity level for each variable (4). This way, DT-FRM enables its users to efficiently solve their initial problem while minimizing DT implementation and recurring costs. It is shown that assessing the appropriate level of DT fidelity is crucial to realize benefits and reduce implementation complexity in manufacturing.

Keywords

Digital Twin; Virtual Twin; Fidelity; Requirements; Benefits; Value; Digital Shadow; Industry 4.0

1. Introduction

Industrial manufacturing is becoming increasingly individual and complex [1]. Organizations must become more agile to satisfy changing customer needs faster and better. In today's globalized economy, they are under constant pressure to improve their performance [2]. One way to meet the increasing competitiveness is digitalization [3]. In the context of Smart Factory and Industry 4.0, there is a wide range of technologies that can be used for this purpose [4], [5]. One of the promising concepts is the Digital Twin (DT). In recent years, the number of scientific publications on the subject has increased exponentially [6]–[8]. At the same time, many companies, especially large corporations, are launching initiatives to explore the potential of DTs [9]–[12]. Despite this attention, the definition of DT remains controversial. The large number of publications has resulted in a multitude of definitions, each with its own specifics [13]–[16]. Their

DOI: https://doi.org/10.15488/12145

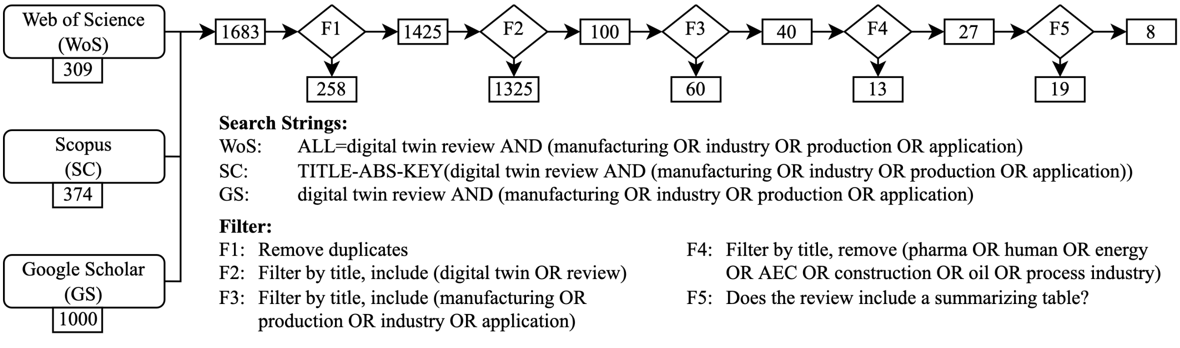
understanding differs significantly, depending on the industry, use case, and context. The most accepted definitions are from [17][1], who first established the Digital Twin concept, and [18][2]. The term DT is frequently used to profit from the hype [19] surrounding the concept. Claimed implementations are often just Digital Models (DM) or Digital Shadows (DS), which, based on [15], merely represent subcategories of DTs. Additionally, the value added by DTs is commonly unclear and intangible [5], [7], [13], [20]. This prevents the unbiased assessment of investments into DT technology and leads to a lack of acceptance within organizations. If organizations are still willing to invest, they often introduce such technology as an end in itself, with no strategy beyond demonstration [21].

One of the reasons why DT economic benefits are difficult to grasp is that the necessary fidelity seems not to be sufficiently considered. According to [7], fidelity indicates "the number of parameters, their accuracy, and level of abstraction". In line with most academic definitions, it is assumed that DTs have to replicate the physical world as realistically as possible, i.e., in high-fidelity [18], [22]–[29]. Thereby, the DT benefits from the rapid technological progress of closely related technologies, such as Simulation, Internet of Things (IoT), Cyber-Physical Systems (CPS), Artificial Intelligence (AI) and Big Data [4], [6], [30]. In practice, however, organizations focus on achieving improvements with minimum effort. Therefore, it is questionable whether it is mandatory to create all-encompassing DTs [7], [31]. In Simulation, which is a core technique of DT [8], [32], focusing on relevant system elements instead of mapping all of its properties, behaviors and states is preferred [33], [34]. In fact, lower fidelity equals less cost compared to high-fidelity [31], [35]. Currently, there is no approach to bridge this gap between academic definitions and practical requirements with regard to DT fidelity. For this reason, the Digital Twin Fidelity Requirements Model (DT-FRM) is presented in this paper. The following sections are structured as follows: Section 2 includes a literary analysis of articles describing DT applications, Section 3 first puts the research question into a broader context and then explains the DT-FRM in detail, and Section 4 summarizes the research findings and discusses implications for scholars and practitioners.

2. Literary Analysis

A literary analysis was conducted to examine current DT literature regarding its implementation procedure, investigating whether the identified articles describe structured implementation procedures. The analysis focused on how fidelity is considered in scholarly described DT applications.

2.1 Methodology



Search Strings:
WoS: ALL=digital twin review AND (manufacturing OR industry OR production OR application)
SC: TITLE-ABS-KEY(digital twin review AND (manufacturing OR industry OR production OR application))
GS: digital twin review AND (manufacturing OR industry OR production OR application)

Filter:
F1: Remove duplicates
F2: Filter by title, include (digital twin OR review)
F3: Filter by title, include (manufacturing OR production OR industry OR application)
F4: Filter by title, remove (pharma OR human OR energy OR AEC OR construction OR oil OR process industry)
F5: Does the review include a summarizing table?

Figure 1: Literature selection process

---

[1] "The Digital Twin concept model […] contains three main parts: a) physical products in Real Space, b) virtual products in Virtual Space, and c) the connections of data and information that ties the virtual and real products together."

[2] "A Digital Twin is an integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its corresponding flying twin. The Digital Twin is ultra-realistic and may consider one or more important and interdependent vehicle systems, including airframe, propulsion and energy storage, life support, avionics, thermal protection, etc."

Since there are an extensive number of publications in the field of DT, a two-step approach for finding relevant articles of actual DT applications was used. First, recent DT review articles were identified as such reviews usually include useful categorizations of DT applications. Second, relevant DT application articles were selected from these reviews. The search strings shown in Figure 1 were used to collect all matching articles from Web of Science, Scopus and Google Scholar. Figure 1 also visualizes the general literature selection process and all applied filters. From Google Scholar, only the first 1000 entries were included. The results of all three search engines were merged into one repository and duplicates were removed (F1). In the following steps, results were further refined by filtering for relevant titles (F2-F4). 27 review articles remained for further analysis. This was done by scanning the articles in question for tables providing structured information of considered DT applications. Finally, eight review articles including such tables were identified. Table 1 illustrates which application articles were chosen from each review for detailed investigation. The column "Selection criteria" refers to the review article's categorization by which application articles were selected. From these reviews, 77 application articles were extracted. They were analyzed in detail to what extent they have considered DT fidelity requirements.

Table 1: Review articles, corresponding application articles and selection criteria

Review articles	Number of articles reviewed	Selected application articles	Selection criteria
[16]	26	[36]–[40]	Manufacturing context
[41]	10	[42]–[46]	Manufacturing context
[15]	43	[32], [35], [42], [47]–[56]	Level of integration DT or DS & type case-study
[57]	32	[47], [58]–[69]	Manufacturing phase
[6]	39	[58]–[60], [70]–[83]	Manufacturing phase
[84]	52	[61], [74], [85]–[98]	Control of real system from DT
[99]	40	[12], [14], [46], [60], [100]–[107]	Application examples (A)
[108]	12	[109]–[112]	Level of integration DT or DS/DT
Sum (duplicates removed)		Σ 85 (77)	

2.2 Results

In summary, it can be confirmed that the understanding of DT among the authors is heterogeneous. Regardless of this, it was first analyzed whether the DT application articles describe a procedure for creating or implementing their DT. Figure 2 shows that 60 articles (78%) do not present any procedure at all. They only describe their individual final solution or architecture, e.g. [58], [60], [68], [74], [80], [85], [90], [98], [110], but not how it has been achieved.

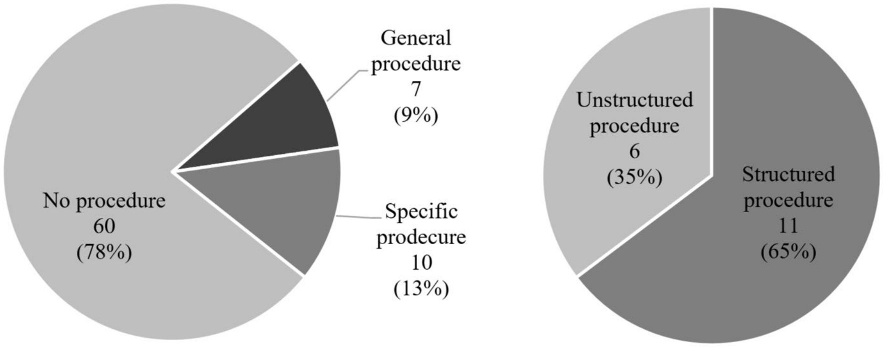


Figure 2: Share of articles describing their DT implementation procedure

The articles with a description of the procedure can be divided into specific [45], [62], [64], [72], [87], [100]–[102], [106], [111] and general procedures [37], [46], [66], [67], [69], [78], [88]. A specific procedure is explicitly tailored to a particular use case, while a general procedure is also transferable to other similar applications. In about two thirds (65%) of the applications described, a structured procedure is recognizable that includes certain steps and sequences. Only one single article from the sample considers fidelity within its procedure. Although the term fidelity is not used directly, an iterative model evolution procedure exists in [67], which adjusts the fidelity step by step to the necessary degree. However, the ultimate goal in [67] is also a high-fidelity model. Due to the low consideration of fidelity within the described procedures, it was investigated whether fidelity is considered in general within the application articles. Figure 3 illustrates the results.

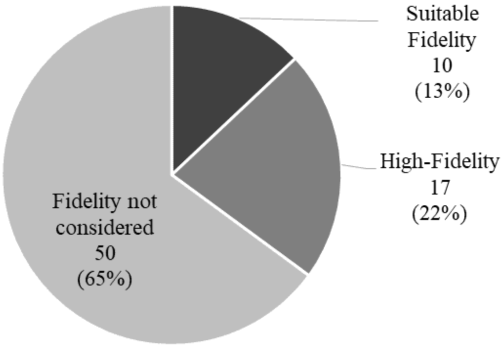


Figure 3: Share of articles considering fidelity

Almost two thirds of the articles do not address fidelity at all. 17 articles (22%) [12], [40], [42], [44], [46], [54], [60], [65], [67], [69], [74], [83], [90], [102]–[105] share the view that DT should represent the physical world in high-fidelity, with most articles referring to the NASA DT definition [18]. The authors usually do not question this definition with regard to fidelity. Only a minority of 10 articles (13%) [14], [32], [35], [37], [55], [56], [62], [97], [101], [110] mention that a suitable fidelity should be chosen. A dominant opinion comes from [32], who clearly mention that an application-specific fidelity should be selected for the DT to achieve a desired goal. [55] cite [32] and adopt their view. Moreover, [14], [37], [56] mention that a specific level of detail should be considered.

Nevertheless, the benefits of applying DTs remain unclear in most articles. In [72] the increase in resource efficiency is evaluated and quantified. However, the authors neglect the cost of implementing the DT and only focus on the positive impact. To achieve economic benefits with the application of DTs, a structured approach must be developed that also takes the necessary DT fidelity level into account since fidelity significantly drives costs.

3. Digital Twin Fidelity

This section first puts the DT-FRM into a broader perspective by highlighting its relevance inside a cost-benefit analysis. Then the methodology for the development of the DT-FRM is explained. Finally, the DT-FRM is presented in detail.

3.1 Cost-Benefit Analysis for Digital Twin Implementation

The decision whether to implement a DT is a complex task. A cost-benefit analysis [21], [113], [114] must be conducted prior to the DT introduction to support an investment decision for or against the use of DT. Figure 4 describes such a procedure for a problem centric cost-benefit analysis based on the DT-FRM. All individual steps and their connections are briefly described below. This section shall help to increase the understanding of how the DT-FRM improves DT implementation decisions.

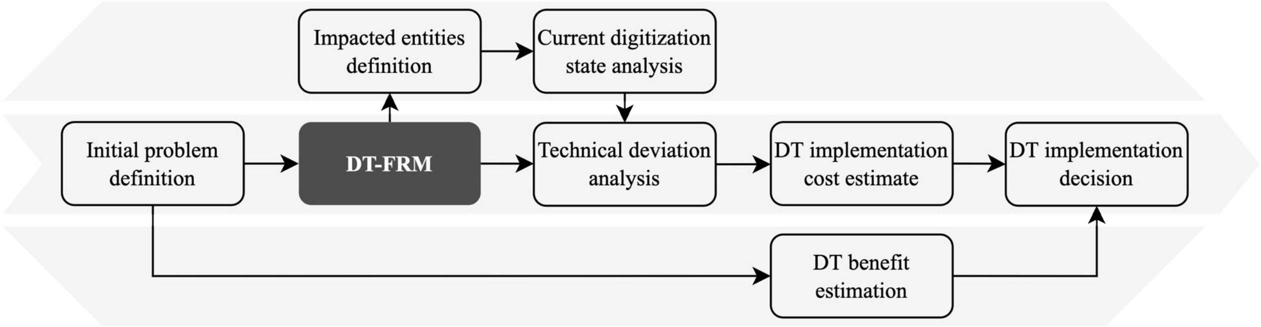


Figure 4: Procedure for Digital Twin cost-benefit analysis

3.1.1 Initial problem definition

First, an existing problem in production must be identified. Once a Digital Twin is perceived to be a viable solution to this problem acknowledged by all stakeholders, an initial problem statement has to be formulated. The problem statement is the foundation of the entire project and facilitates common understanding within the project team. It should therefore precisely describe what constitutes a problem in the current state of a production.

3.1.2 Digital Twin Fidelity Requirements Model (DT-FRM)

This article’s main contribution is a structured approach for the elaboration of the required fidelity level for a specific DT implementation serving as a solution to the initially described problem. The DT-FRM is presented in detail in section 3.2.

3.1.3 Impacted entities definition

Following the joint agreement on the problem to be solved by DT implementation, an investigation is needed to identify impacted entities. In manufacturing environments, these might be products, processes and resources. Every relevant variable identified in the DT-FRM has to correspond to at least one entity. While reviewing those entities, the focus always needs to be on the initial problem.

3.1.4 Current digitization state analysis

After identifying all entities which are impacted by the initially defined problem, a technological analysis must be carried out. The result of such an analysis is a detailed overview of the current state of digitization, e.g., data, model or control loop availability. It includes an assessment of all relevant entities for a subsequent estimate of the technical changes required to introduce a DT.

3.1.5 Technical deviation analysis

For each identified entity, the individual deviation between the required fidelity level, which has been elaborated within the DT-FRM, and the current state of digitization has to be determined. Some variables might already be digitally mapped or even controlled autonomously in current production. For this reason, it is necessary to identify gaps while working towards the elaborated level of fidelity.

3.1.6 Digital Twin implementation cost estimate

After the necessary changes in production have been identified in the technical deviation analysis, they must now be evaluated in terms of additional cost. This is where the individual comparison between actual digitization state and required DT fidelity becomes important. The identified deltas give guidance for estimating recurring and non-recurring costs to reach the desired target state and achieve initial problem solution.

3.1.7 Digital Twin benefit estimation

Based on the initial problem defined, benefits have to be estimated. The calculation is carried out independently of a specific technical implementation and its costs, purely on the basis of potential savings achieved by a still undefined solution. The aim of this analysis is to make an initial statement about the savings that can be expected as a result of fully solving the central problem.

3.1.8 Digital Twin implementation decision

With necessary changes for DT implementation evaluated financially, a final decision on DT implementation must be taken. Therefore, estimated benefits of solving the initial problem are directly compared with estimated costs of achieving required fidelity levels. The present value of all cash flows must be calculated for determining the overall net present value (NPV).

3.2 Development of Digital Twin Fidelity Requirements Model (DT-FRM)

The DT-FRM has been developed by employing concepts taken from the Design Science Research (DSR) methodology [115]. DSR has become the leading approach in the development of information systems [116]. Since the technologies around DTs are based on such systems, applying DSR seems adequate. Design Science comprises two iterative activities: the design cycle and the empirical cycle, which are used for the design and investigation of artifacts in different contexts [116]. Artifacts are single solutions to a problem within a specific context. Using the DSR template from [116], the research question for design is formulated as “How to develop a method that considers appropriate Digital Twin fidelity requirements so that users can increase the likelihood of achieving economic benefits by implementing Digital Twins in manufacturing to solve existing problems?” In this case, the DT-FRM is the final artifact resulting from several iterations of the design cycle. For the development of the DT-FRM, only the design cycle was needed. It includes three steps: problem investigation, treatment design and treatment validation [116]. Here, treatment refers to the desired interaction of artifact and problem context. The DT-FRM is designed as a universal artifact which can be applied to different contexts, i.e. DT application scenarios. Knowledge questions then have to be answered around this specific context. Whenever it is intended to implement DTs in manufacturing, DT-FRM can be used to assist with problem centric fidelity assessment. In the DSR design cycle, validation occurs before implementation and is done by predicting the artifacts’ behavior within a given context [116].

3.3 Digital Twin Fidelity Requirements Model (DT-FRM)

This section presents a structured method for the elaboration of fidelity requirements for DTs in production environments, called DT-FRM. Employing the DT-FRM is a crucial part of the cost assessment within the cost-benefit analysis as higher fidelity is associated with higher costs. Therefore, considering appropriate fidelity levels contributes to achieving economic benefits when applying DTs for problem solving. Based on these requirements, an implementation strategy can be derived that provides an efficient solution to the problem statement described initially.

3.3.1 Target variable identification (TV)

The DT-FRM focuses on the decomposition of the initial problem (Section 3.1.1) into its quantifiable components. Therefore, the first step of the DT-FRM is to define target variables (TVs) which are often called Key Performance Indicators (KPIs) in practice. TVs are usually a set of KPIs which are already regularly calculated for monitoring manufacturing performance. Independent of the number of TVs a problem is represented by, the desired direction and magnitude of change for each variable towards the problem solution must be defined. In an example, a defined problem might be characterized primarily by one single KPI. Then for this TV, it needs to be determined whether an increase or a decrease contributes to solving the initial problem and how much the value must change. In a problem graph (Figure 5), the TVs represent the first layer. They are called dependent variables, since their value is dependent on a variety of other, underlying variables.

3.3.2 Intermediate (IV) and elementary variable (EV) derivation

To ensure that the initial problem is comprehensively broken down into its relevant and, in particular, influenceable components, the derivation of the TVs must be followed by a detailed examination of their calculation basis. This has to be done for each KPI defined as a TV in the previous step. If, for example, the initial problem is from the field of machining, a possible TV could be the tool life T. Typically, the tool life results from the theoretical tool life c_v, the cutting speed v_c and the slope of the Taylor line κ. The TV tool life is thus dependent on these three underlying variables, which are referred to as intermediate variables (IVs) in the DT-FRM. IVs neither serve as a reference to the initial problem, nor can they directly be influenced. For complex problems in real manufacturing environments, it is common that the derivation of IVs yields multiple layers of interlaced variables. The goal of the decomposition of TVs into their calculation basis (IV) is the elaboration of all directly influenceable, fundamental variables. These variables are called elementary variables (EV) in the DT-FRM context. They are not based on any underlying variables and are, therefore, independent. In the simple example of tool life as a TV, an EV is the rotational speed of a machine, which in turn has an effect on the cutting speed (IV) of the machining operation. The EV rotational speed in this example can be considered as independent and therefore directly influenced by applying DT technology. Finally, an overall picture of the initial problem and its influenceable variables is obtained: all identified EVs ultimately result in the TVs defined at the beginning by calculating all IVs. Figure 5 illustrates the dependencies of TVs, IVs and EVs for a schematic problem. The use of such problem graphs in complex manufacturing scenarios is especially helpful to identify overlapping influences of individual variables and to provide a uniform understanding among all stakeholders.

3.3.3 Elementary variable (EV) prioritization

The goal of this step is the prioritization of EVs. All EVs must be evaluated according to their influenceability and their target contribution. For determining the influenceability, an optimization corridor around the current mean value must be defined for each EV. The optimization corridor determines to what extent a change in the corresponding EV is estimated to be realistically achievable, based on financial, technical or organizational constraints. Financial constraints refer to the costs of influencing the EV, technical constraints refer to technological feasibility and organizational constraints are based on the structure of the organization aiming to apply DT solutions. Since estimating the boundaries of the optimization corridor and defining the mathematical relationships between the variables are highly case-specific, a certain experience in the problem context is necessary. If this knowledge is not available within

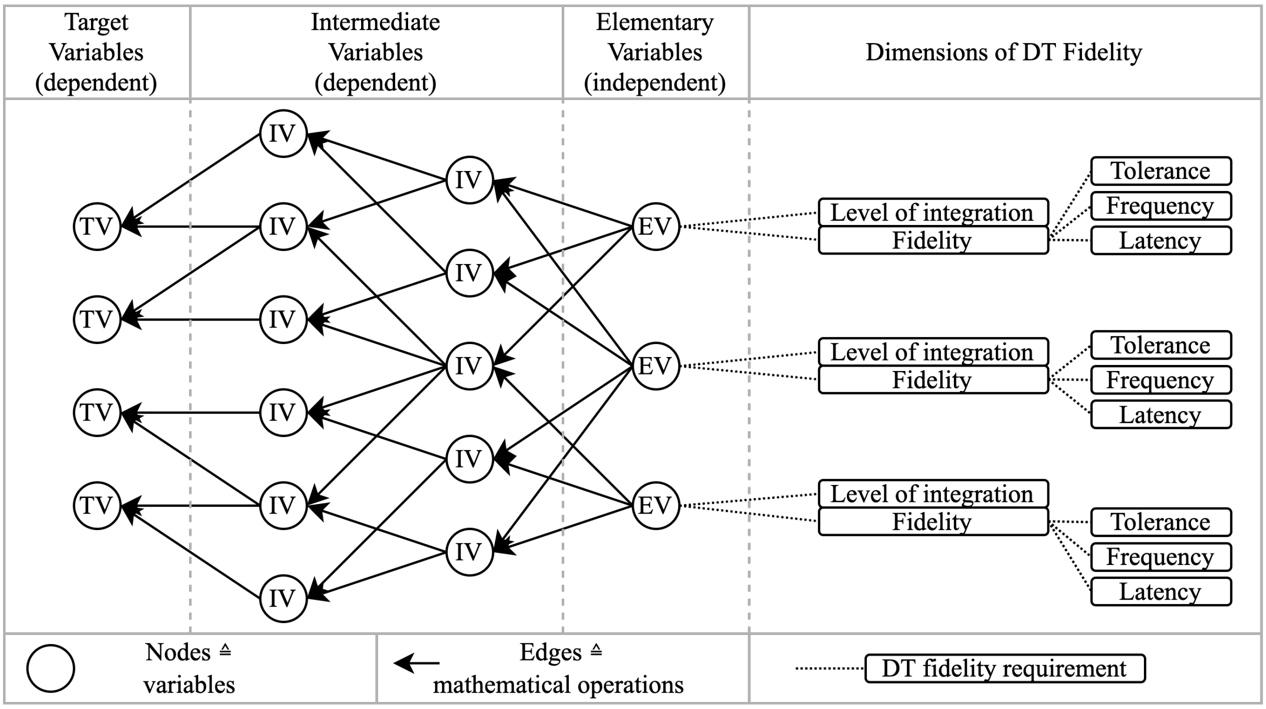


Figure 5: Example problem graph with variable breakdown

the organization, it has to be questioned whether applying DT solutions is effective. Before implementing DTs, it is necessary to clearly understand the initial problem and the implementation objectives (Section 3.1.1). The second step of the prioritization involves conducting a sensitivity analysis to identify target contribution. By conducting a sensitivity analysis, the potential impact of each EV change towards the problem solution is determined. Minimum and maximum values of the optimization corridor serve as input for sensitivity analysis. Ultimately, the EVs which provide the highest influenceability and highest target contribution are prioritized within the next steps to minimize required efforts and maximize benefits. EVs with low influenceability or low target contribution can be neglected in a first step. Figure 6 illustrates a matrix for EV prioritization with different sectors and respective priorities.

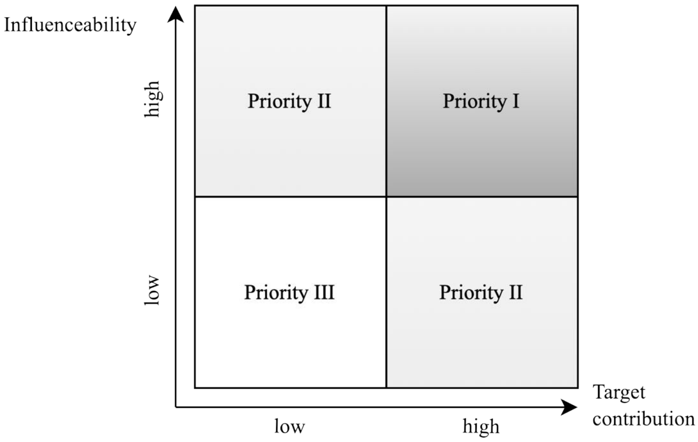


Figure 6: Elementary variables priority matrix

3.3.4 Elementary variable (EV) fidelity elaboration

Once the initial problem is broken down into its underlying EVs, the actual DT concept has to be developed. DT is commonly considered as an ultra-realistic, high-fidelity virtual model of a physical entity, mirroring all of its properties most accurately [13].

Supporting the understanding of what characterizes a comprehensive DT, we concurrently question whether all EVs actually have to receive such treatment in reality. The DT fidelity required for a variable to support target contribution is highly context dependent. Therefore, the DT-FRM proposes a different approach to put DT technology into beneficial use in industrial applications. Starting with priority I EVs, the individual variables are assessed for their required fidelity to support TV adjustment towards problem solution. Generally sharing the understanding of [7] in terms of fidelity, the DT-FRM introduces two overall dimensions which are used to determine DT fidelity requirements: level of integration (1) and fidelity (2). Since the meaning of concepts around fidelity like abstraction, accuracy, granularity, precision, etc., is similar but not identical [117], this article additionally defines three sub-dimensions of fidelity.

The first overall dimension is the level of integration. [15] defines the three different DT levels of integration: modeling, shadowing and twinning. Modeling refers to manual data exchange from physical to virtual (P2V) and virtual to physical (V2P). Shadowing describes P2V as fully automatic with V2P still being manual. Twinning is then understood as automatic P2V with the feedback loop V2P being automatic as well. For every EV incorporated into a DT application, the level of integration has to be defined. If the variable needs to be monitored autonomously and digital control is required to alter its value in terms of target contribution, the level of integration to implement is twinning. If monitoring is required but no automatic control is needed, the level of integration is shadowing. If none is the case, modeling is sufficient for the particular variable. The second overall dimension is fidelity, which consists of three sub-dimensions: tolerance (1), frequency (2) and latency (3). For each EV, the technical tolerance for measuring and, in the case of twinning, for control needs to be determined. The tolerance defines how precisely a value needs to be monitored or altered to achieve target contribution. Furthermore, the frequency needed for data exchange between the DTs physical and virtual space needs to be considered. Frequency thereby is regarded as how often data is transferred during a given time interval. The third sub-dimension is latency. Latency describes the amount of time data needs to reach its destination, which is also known as delay. Instead of using scarce financial resources to reach out for maximum fidelity, it must be carefully evaluated which minimum level is required to secure the respective variables' target contribution. Otherwise, over-engineering fidelity leads to excess costs, which must be avoided. Thus, not all EVs require high-fidelity twinning. After the level of integration and the DT fidelity are elaborated for all relevant variables, the EVs can be numbered and plotted into a DT fidelity requirements matrix. Figure 7 gives a basic example of such a matrix. By utilizing such matrices, the overall complexity of proposed DT solutions to different problems can be visualized. The higher the level of integration and fidelity, the higher the estimated costs for implementing the DT.

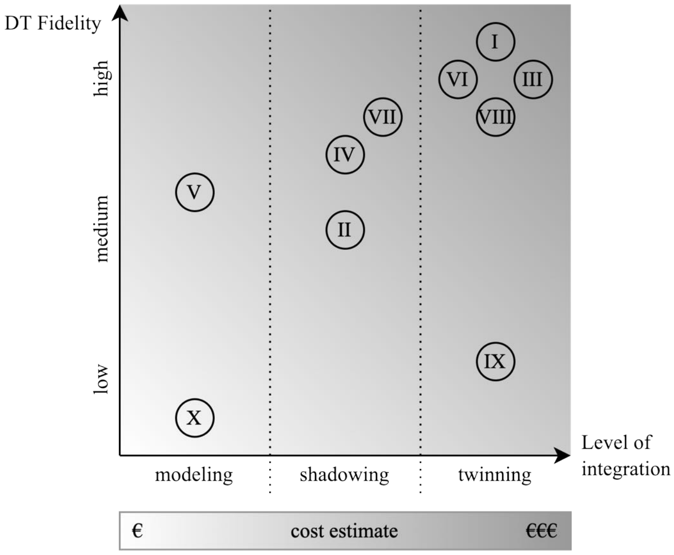


Figure 7: Digital Twin fidelity requirements matrix with example elementary variables

4. Results and Discussion

In today's complex manufacturing environments, organizations face highly competitive pressure and are therefore dependent on promising digitalization concepts like Digital Twins. However, the understanding of what a DT is and how it can effectively be applied to solve existing problems differs among organizations. According to most academic definitions, the fidelity of the virtual models replicating the physical world must be as high as possible, whereas in practice this is not always feasible. Literary analysis of 77 scientific papers describing DT applications in manufacturing has revealed a lack of conceptual basis and guidelines for structured implementation of DTs. This hinders the applicability of DTs in different domains. Even if structured procedures have been described, they tend to be specific and not transferable. Additionally, DT fidelity has not been considered in most application articles. The majority have been found to support academia's common understanding of targeting high-fidelity. Contrary, this article presented the Digital Twin Fidelity Requirements Model (DT-FRM) as part of a cost-benefit analysis for DT implementation decisions. The DT-FRM aims at securing economic benefits when applying DT technology to exploit existing improvement potentials in production environments. Despite questioning the focus of most academics aiming for high-fidelity models, we do not generally reject the available definitions of DT. Instead, we emphasize that elaborating suitable fidelity levels is necessary to maximize benefits by applying DTs to existing problems. Since concrete benefits of using DTs are currently still unclear, applying the DT-FRM to defined problems serves as a good starting point to increase understanding and decrease implementation complexity of DTs and its related technologies. The method helps practitioners to estimate benefits of DT application while assisting with DT concept development. Nevertheless, we suppose that the benefits of applying DTs in the future go beyond merely solving known problems, e.g., by unveiling hidden improvement potentials and enabling new business models. Iteratively increasing fidelity during the lifetime of the DT might be a solution to exploit its full potential while still considering appropriate fidelity levels. Future research should address the application of the DT-FRM to real manufacturing scenarios to confirm its necessity and validity. Additionally, it should be investigated which other factors besides fidelity influence costs for DT implementation.

Acknowledgements

This research is funded by dtec.bw within the project "LaiLa - Laboratory for intelligent lightweight applications". We would like to thank the Composite Technology Center / CTC GmbH (An Airbus Company) for supporting our work.

References

[1] Bauernhansl, T., Krüger, J., Reinhart, G., Schuh, G., 2016. WGP-Standpunkt Industrie 4.0. Wissenschaftliche Gesellschaft für Produktionstechnik e.V. – WGP, Accessed: 07/30/2021, [Online], Available: https://wgp.de/wp-content/uploads/WGP-Standpunkt_Industrie_4-0.pdf

[2] Sanders, A., Elangeswaran, C., Wulfsberg, J., 2016. Industry 4.0 implies lean manufacturing: Research activities in industry 4.0 function as enablers for lean manufacturing. Journal of Industrial Engineering and Management 9 (3), 811.

[3] Björkdahl, J., 2020. Strategies for Digitalization in Manufacturing Firms. California Management Review 62 (4), 17–36.

[4] Oztemel, E., Gursev, S., 2020. Literature review of Industry 4.0 and related technologies. Journal of Intelligent Manufacturing 31 (1), 127–182.

[5] Osterrieder, P., Budde, L., Friedli, T., 2020. The smart factory as a key construct of industry 4.0: A systematic literature review. International Journal of Production Economics 221.

[6] Liu, M., Shuiliang, F., Dong, H., Xu, C., 2020. Review of digital twin about concepts, technologies, and industrial applications. Journal of Manufacturing Systems 58 (Part B), 346–361.

[7] Jones, D., Snider, C., Nassehi, A., Yon, J., Hicks, B., 2020. Characterising the Digital Twin: A systematic literature review. CIRP Journal of Manufacturing Science and Technology 29, 36–52.

[8] Tao, F., Zhang, H., Liu, A., Nee, A.Y.C., 2019. Digital Twin in Industry: State-of-the-Art. IEEE Transactions on Industrial Informatics 15 (4), 2405–2415.

[9] Wendenburg, M., 2017. Digital Twin is about to rollout by Airbus. Accessed: 01/27/2022, [Online], Available: https://ascon-systems.de/en/digital-twin-is-about-to-rollout-by-airbus/

[10] Daimler, Production is becoming smart. Industry 4.0 and the networked factory. Accessed: 01/27/2022, [Online], Available: https://www.daimler.com/innovation/case/connectivity/industry-4-0.html

[11] Caulfield, B., 2021. NVIDIA, BMW Blend Reality, Virtual Worlds to Demonstrate Factory of the Future. Accessed: 01/27/2022, [Online], Available: https://blogs.nvidia.com/blog/2021/04/13/nvidia-bmw-factory-future/

[12] General Electric, 2016. GE Digital Twin: Analytic Engine for the Digital Power Plant. Accessed: 01/08/2022, [Online], Available: https://www.ge.com/digital/sites/default/files/download_assets/Digital-Twin-for-the-digital-power-plant-.pdf

[13] VanDerHorn, E., Mahadevan, S., 2021. Digital Twin: Generalization, characterization and implementation. Decision Support Systems 145.

[14] Uhlenkamp, J.-F., Hribernik, K., Wellsandt, S., Thoben, K.-D., 2019. Digital Twin Applications : A first systemization of their dimensions. IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Valbonne Sophia-Antipolis, France, 1–8.

[15] Kritzinger, W., Karner, M., Traar, G., Henjes, J., Sihn, W., 2018. Digital Twin in manufacturing: A categorical literature review and classification. IFAC-PapersOnLine 51 (11), 1016–1022.

[16] Negri, E., Fumagalli, L., Macchi, M., 2017. A Review of the Roles of Digital Twin in CPS-based Production Systems. Procedia Manufacturing 11, 939–948.

[17] Grieves, M., 2015. Digital Twin: Manufacturing Excellence through Virtual Factory Replication. Accessed: 01/08/2022, [Online], Available: https://www.3ds.com/fileadmin/PRODUCTS-SERVICES/DELMIA/PDF/Whitepaper/DELMIA-APRISO-Digital-Twin-Whitepaper.pdf

[18] Glaessgen, E.H., Stargel, D.S., 2012. The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. Accessed: 01/08/2022, [Online], Available: https://ntrs.nasa.gov/citations/20120008178

[19] Panetta, K., 2019. 5 Trends Emerge In Gartner Hype Cycle For Emerging Technologies 2018. Accessed: 01/27/2022, [Online], Available: https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018

[20] Polini, W., Corrado, A., 2020. Digital twin of composite assembly manufacturing process. International Journal of Production Research 58 (17), 5238–5252.

[21] Joppen, R., Lipsmeier, A., Tewes, C., Kühn, A., Dumitrescu, R., 2019. Evaluation of investments in the digitalization of a production. Procedia CIRP 81, 411–416.

[22] Reifsnider, K., Majumdar, P., 2013. Multiphysics Stimulated Simulation Digital Twin Methods for Fleet Management. 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Boston, USA.

[23] Shafto, M., Conroy, M., Doyle, R., Glaessgen, E.H., Kemp, C., LeMoigne, J., Wang, L., 2010. Modeling, Simulation, Information Technology & Processing Roadmap Technology Area 11. Accessed: 01/27/2022, [Online], Available: https://www.nasa.gov/pdf/501321main_TA11-MSITP-DRAFT-Nov2010-A1.pdf

[24] Bielefeldt, B.R., Hochhalter, J.D., Hartl, D.J., 2018. Shape memory alloy sensory particles for damage detection: Experiments, analysis, and design studies. Structural Health Monitoring 17 (4), 777–814.

[25] Bazilevs, Y., Deng, X., Korobenko, A., Lanza di Scalea, F., Todd, M.D., Taylor, S.G., 2015. Isogeometric Fatigue Damage Prediction in Large-Scale Composite Structures Driven by Dynamic Sensor Data. Journal of Applied Mechanics 82 (9).

[26] Grieves, M., Vickers, J., 2017. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. in: Kahlen, F.-J., Flumerfelt, S., Alves, A. (Eds.), Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, Springer International Publishing, Cham, pp. 85–113.

[27] Alam, K.M., El Saddik, A., 2017. C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. IEEE Access 5, 2050–2062.

[28] Zheng, P., Lin, T.-J., Chen, C.-H., Xu, X., 2018. A systematic design approach for service innovation of smart product-service systems. Journal of Cleaner Production 201, 657–667.

[29] Talkhestani, B.A., Jazdi, N., Schloegl, W., Weyrich, M., 2018. Consistency check to synchronize the Digital Twin of manufacturing automation based on anchor points. Procedia CIRP 72, 159–164.

[30] Tao, F., Qi, Q., Wang, L., Nee, A.Y.C., 2019. Digital Twins and Cyber–Physical Systems toward Smart Manufacturing and Industry 4.0: Correlation and Comparison. Engineering 5 (4), 653–661.

[31] Zhang, L., Zhou, L., Horn, B.K.P., 2021. Building a right digital twin with model engineering. Journal of Manufacturing Systems 59, 151–164.

[32] Boschert, S., Rosen, R., 2016. Digital Twin—The Simulation Aspect. in: Hehenberger, P., Bradley, D. (Eds.), Mechatronic Futures, Springer International Publishing, Cham, pp. 59–74.

[33] van der Zee, D.-J., 2019. Model simplification in manufacturing simulation – Review and framework. Computers & Industrial Engineering 127, 1056–1067.

[34] Fishwick, P.A., 1988. The role of process abstraction in simulation. IEEE Transactions on Systems, Man, and Cybernetics 18 (1), 18–39.

[35] Müller, R., Vette, M., Hörauf, L., Speicher, C., Burkhard, D., 2017. Lean Information and Communication Tool to Connect Shop and Top Floor in Small and Medium-sized Enterprises. Procedia Manufacturing 11, 1043–1052.

[36] Arisoy, E.B., Ren, G., Ulu, E., Ulu, N.G., Musuvathy, S., 2016. A Data-Driven Approach to Predict Hand Positions for Two-Hand Grasps of Industrial Objects. 36th Computers and Information in Engineering Conference, Charlotte, USA.

[37] Schroeder, G.N., Steinmetz, C., Pereira, C.E., Espindola, D.B., 2016. Digital Twin Data Modeling with AutomationML and a Communication Methodology for Data Exchange. IFAC-PapersOnLine 49 (30), 12–17.

[38] Abramovici, M., Göbel, J.C., Dang, H.B., 2016. Semantic data management for the development and continuous reconfiguration of smart products and systems. CIRP Annals 65 (1), 185–188.

[39] Rosen, R., von Wichert, G., Lo, G., Bettenhausen, K.D., 2015. About The Importance of Autonomy and Digital Twins for the Future of Manufacturing. IFAC-PapersOnLine 48 (3), 567–572.

[40] Gabor, T., Belzner, L., Kiermeier, M., Beck, M.T., Neitz, A., 2016. A Simulation-Based Architecture for Smart Cyber-Physical Systems. IEEE International Conference on Autonomic Computing (ICAC), Wuerzburg, Germany, 374–379.

[41] Campos-Ferreira, A., Lozoya-Santos, J., Vargas-Martınez, A., Mendoza, R.R., Morales-Menendez, R., 2019. Digital Twin Applications: A Review. Memorias del congreso nacional de control automático, Puebla, Mexico.

[42] Vachalek, J., Bartalsky, L., Rovny, O., Sismisova, D., Morhac, M., Loksik, M., 2017. The digital twin of an industrial production line within the industry 4.0 concept. 21st International Conference on Process Control, Strbske Pleso, Slovakia, 258–262.

[43] Rodič, B., 2017. Industry 4.0 and the New Simulation Modelling Paradigm. Organizacija 50 (3), 193–207.

[44] Haag, S., Anderl, R., 2018. Digital twin – Proof of concept. Manufacturing Letters 15, 64–66.

[45] Luo, W., Hu, T., Zhu, W., Tao, F., 2018. Digital twin modeling method for CNC machine tool. IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, China, 1–4.

[46] Parrott, A., Warshaw, L., 2017. Industry 4.0 and the digital twin. [Online], Available: https://www2.deloitte.com/us/en/insights/focus/industry-4-0/digital-twin-technology-smart-factory.html

[47] Uhlemann, T.H.-J., Lehmann, C., Steinhilper, R., 2017. The Digital Twin: Realizing the Cyber-Physical Production System for Industry 4.0. Procedia CIRP 61, 335–340.

[48] Bottani, E., Cammardella, A., Murino, T., Vespoli, S., 2017. From the Cyber-Physical System to the Digital Twin: the process development for behaviour modelling of a Cyber Guided Vehicle in M2M logic. XXII Summer School F. Turco - Industrial Systems Engineering, 1–7.

[49] Brenner, B., Hummel, V., 2017. Digital Twin as Enabler for an Innovative Digital Shopfloor Management System in the ESB Logistics Learning Factory at Reutlingen - University. Procedia Manufacturing 9, 198–205.

[50] Kuhn, T., 2017. Digitaler Zwilling. Informatik-Spektrum 40 (5), 440–444.

[51] Lindström, J., Larsson, H., Jonsson, M., Lejon, E., 2017. Towards Intelligent and Sustainable Production: Combining and Integrating Online Predictive Maintenance and Continuous Quality Control. Procedia CIRP 63, 443–448.

[52] Mell, P., Grance, T., 2011. The NIST definition of cloud computing. NIST SP 800-145, National Institute of Standards and Technology, Accessed: 01/10/2022, [Online], Available: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf

[53] Pawlaszczyk, D., 2006. Scalable multi agent based simulation-Considering efficient simulation of transport logistic networks. 12th ASIM Conference - Simulation in Production and Logistics, Kassel, Germany.

[54] Reeves, C., 2017. Spotlight on the Digital Twin. [Online], Available: https://www.ansys.com/content/dam/product/systems-embedded-and-integrated/twin-builder/ansys-advantage-digital-twin-aa-v11-i1.pdf

[55] Schleich, B., Anwer, N., Mathieu, L., Wartzack, S., 2017. Shaping the digital twin for design and production engineering. CIRP Annals 66 (1), 141–144.

[56] Schluse, M., Rossmann, J., 2016. From simulation to experimentable digital twins: Simulation-based development and operation of complex technical systems. IEEE International Symposium on Systems Engineering (ISSE), Edinburgh, United Kingdom, 1–6.

[57] Hu, W., Zhang, T., Deng, X., Liu, Z., Tan, J., 2021. Digital twin: a state-of-the-art review of its enabling technologies, applications and challenges. Journal of Intelligent Manufacturing and Special Equipment 2 (1), 1–34.

[58] Leng, J., Zhang, H., Yan, D., Liu, Q., Chen, X., Zhang, D., 2019. Digital twin-driven manufacturing cyber-physical system for parallel controlling of smart workshop. Journal of Ambient Intelligence and Humanized Computing 10 (3), 1155–1166.

[59] Malik, A.A., Bilberg, A., 2018. Digital twins of human robot collaboration in a production setting. Procedia Manufacturing 17, 278–285.

[60] Zhuang, C., Liu, J., Xiong, H., 2018. Digital twin-based smart production management and control framework for the complex product assembly shop-floor. The International Journal of Advanced Manufacturing Technology 96 (1–4), 1149–1163.

[61] Coronado, P.D.U., Lynn, R., Louhichi, W., Parto, M., Wescoat, E., Kurfess, T., 2018. Part data integration in the Shop Floor Digital Twin: Mobile and cloud technologies to enable a manufacturing execution system. Journal of Manufacturing Systems 48, 25–33.

[62] Liu, Q., Zhang, H., Leng, J., Chen, X., 2019. Digital twin-driven rapid individualised designing of automated flow-shop manufacturing system. International Journal of Production Research 57 (12), 3903–3919.

[63] Priggemeyer, M., Roßmann, J., 2018. Simulation-based Control of Reconfigurable Robotic Workcells: Interactive Planning and Execution of Processes in Cyber-Physical Systems. 50th International Symposium on Robotics, Munich, Germany.

[64] Qamsane, Y., Chen, C.-Y., Balta, E.C., Kao, B.-C., Mohan, S., Moyne, J., Tilbury, D., Barton, K., 2019. A Unified Digital Twin Framework for Real-time Monitoring and Evaluation of Smart Manufacturing Systems. IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, Canada, 1394–1401.

[65] Qi, Q., Tao, F., Zuo, Y., Zhao, D., 2018. Digital Twin Service towards Smart Manufacturing. Procedia CIRP 72, 237–242.

[66] Stark, R., Kind, S., Neumeyer, S., 2017. Innovations in digital modelling for next generation manufacturing system design. CIRP Annals 66 (1), 169–172.

[67] Tao, F., Zhang, M., 2017. Digital Twin Shop-Floor: A New Shop-Floor Paradigm Towards Smart Manufacturing. IEEE Access 5, 20418–20427.

[68] Zambal, S., Eitzinger, C., Clarke, M., Klintworth, J., Mechin, P.-Y., 2018. A digital twin for composite parts manufacturing: Effects of defects analysis based on manufacturing data. IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, 803–808.

[69] Zhang, H., Zhang, G., Yan, Q., 2019. Digital twin-driven cyber-physical production system towards smart shop-floor. Journal of Ambient Intelligence and Humanized Computing 10 (11), 4439–4453.

[70] Bohlin, R., Hagmar, J., Bengtsson, K., Lindkvist, L., Carlson, J.S., Söderberg, R., 2017. Data Flow and Communication Framework Supporting Digital Twin for Geometry Assurance. ASME International Mechanical Engineering Congress and Exposition, Tampa, USA.

[71] Ghosh, A.K., Ullah, A.S., Kubo, A., 2019. Hidden Markov model-based digital twin construction for futuristic manufacturing systems. Artificial Intelligence for Engineering Design, Analysis and Manufacturing 33 (3), 317–331.

[72] Kannan, K., Arunachalam, N., 2019. A Digital Twin for Grinding Wheel: An Information Sharing Platform for Sustainable Grinding Process. Journal of Manufacturing Science and Engineering 141 (2).

[73] Knapp, G.L., Mukherjee, T., Zuback, J.S., Wei, H.L., Palmer, T.A., De, A., DebRoy, T., 2017. Building blocks for a digital twin of additive manufacturing. Acta Materialia 135, 390–399.

[74] Liu, J., Zhou, H., Liu, X., Tian, G., Wu, M., Cao, L., Wang, W., 2019. Dynamic Evaluation Method of Machining Process Planning Based on Digital Twin. IEEE Access 7, 19312–19323.

[75] Macchi, M., Roda, I., Negri, E., Fumagalli, L., 2018. Exploring the role of Digital Twin for Asset Lifecycle Management. IFAC-PapersOnLine 51 (11), 790–795.

[76] Nikolakis, N., Alexopoulos, K., Xanthakis, E., Chryssolouris, G., 2019. The digital twin implementation for linking the virtual representation of human-based production tasks to their physical counterpart in the factory-floor. International Journal of Computer Integrated Manufacturing 32 (1), 1–12.

[77] Pereverzev, P.P., Akintseva, A.V., Alsigar, M.K., Ardashev, D.V., 2019. Designing optimal automatic cycles of round grinding based on the synthesis of digital twin technologies and dynamic programming method. Mechanical Sciences 10 (1), 331–341.

[78] Sharif Ullah, A., 2019. Modeling and simulation of complex manufacturing phenomena using sensor signals from the perspective of Industry 4.0. Advanced Engineering Informatics 39, 1–13.

[79] Soares, R.M., Câmara, M.M., Feital, T., Pinto, J.C., 2019. Digital Twin for Monitoring of Industrial Multi-Effect Evaporation. Processes 7 (8).

[80] Sun, X., Bao, J., Li, J., Zhang, Y., Liu, S., Zhou, B., 2020. A digital twin-driven approach for the assembly-commissioning of high precision products. Robotics and Computer-Integrated Manufacturing 61.

[81] Zhang, H., Zhang, G., Yan, Q., 2018. Dynamic resource allocation optimization for digital twin-driven smart shopfloor. IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), Zhuhai, 1–5.

[82] Zhao, R., Yan, D., Liu, Q., Leng, J., Wan, J., Chen, X., Zhang, X., 2019. Digital Twin-Driven Cyber-Physical System for Autonomously Controlling of Micro Punching System. IEEE Access 7, 9459–9469.

[83] Zhu, Z., Liu, C., Xu, X., 2019. Visualisation of the Digital Twin data in manufacturing by using Augmented Reality. Procedia CIRP 81, 898–903.

[84] Cimino, C., Negri, E., Fumagalli, L., 2019. Review of digital twin applications in manufacturing. Computers in Industry 113.

[85] Angrish, A., Starly, B., Lee, Y.-S., Cohen, P.H., 2017. A flexible data schema and system architecture for the virtualization of manufacturing machines (VMM). Journal of Manufacturing Systems 45, 236–247.

[86] Ardanza, A., Moreno, A., Segura, Á., de la Cruz, M., Aguinaga, D., 2019. Sustainable and flexible industrial human machine interfaces to support adaptable applications in the Industry 4.0 paradigm. International Journal of Production Research 57 (12), 4045–4059.

[87] Moreno, A., Velez, G., Ardanza, A., Barandiaran, I., de Infante, Á.R., Chopitea, R., 2017. Virtualisation process of a sheet metal punching machine within the Industry 4.0 vision. International Journal on Interactive Design and Manufacturing (IJIDeM) 11 (2), 365–373.

[88] Guo, F., Zou, F., Liu, J., Wang, Z., 2018. Working mode in aircraft manufacturing based on digital coordination model. The International Journal of Advanced Manufacturing Technology 98, 1547–1571.

[89] Longo, F., Nicoletti, L., Padovano, A., 2019. Ubiquitous knowledge empowers the Smart Factory: The impacts of a Service-oriented Digital Twin on enterprises' performance. Annual Reviews in Control 47, 221–236.

[90] Zhang, H., Liu, Q., Chen, X., Zhang, D., Leng, J., 2017. A Digital Twin-Based Approach for Designing and Multi-Objective Optimization of Hollow Glass Production Line. IEEE Access 5, 26901–26911.

[91] Um, J., Popper, J., Ruskowski, M., 2018. Modular augmented reality platform for smart operator in production environment. IEEE Industrial Cyber-Physical Systems (ICPS), St. Petersburg, 720–725.

[92] Oyekan, J.O., Hutabarat, W., Tiwari, A., Grech, R., Aung, M.H., Mariani, M.P., López-Dávalos, L., Ricaud, T., Singh, S., Dupuis, C., 2019. The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans. Robotics and Computer-Integrated Manufacturing 55, 41–54.

[93] Park, K.T., Nam, Y.W., Lee, H.S., Im, S.J., Noh, S.D., Son, J.Y., Kim, H., 2019. Design and implementation of a digital twin application for a connected micro smart factory. International Journal of Computer Integrated Manufacturing 32 (6), 596–614.

[94] Hu, L., Nguyen, N.-T., Tao, W., Leu, M.C., Liu, X.F., Shahriar, M.R., Al Sunny, S.M.N., 2018. Modeling of Cloud-Based Digital Twins for Smart Manufacturing with MT Connect. Procedia Manufacturing 26, 1193–1203.

[95] Shahriar, M.R., Sunny, S.M.N.A., Liu, X., Leu, M.C., Hu, L., Nguyen, N.-T., 2018. MTComm Based Virtualization and Integration of Physical Machine Operations with Digital-Twins in Cyber-Physical Manufacturing Cloud. 5th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud) and 4th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom), Shanghai, China, 46–51.

[96] DebRoy, T., Zhang, W., Turner, J., Babu, S.S., 2017. Building digital twins of 3D printing machines. Scripta Materialia 135, 119–124.

[97] Kuts, V., Modoni, G.E., Terkaj, W., Tähemaa, T., Sacco, M., Otto, T., 2017. Exploiting Factory Telemetry to Support Virtual Reality Simulation in Robotics Cell. in: De Paolis, L.T., Bourdot, P., Mongelli, A. (Eds.), Augmented Reality, Virtual Reality, and Computer Graphics, Springer International Publishing, Cham, pp. 212–221.

[98] Souza, V., Cruz, R., Silva, W., Lins, S., Lucena, V., 2019. A Digital Twin Architecture Based on the Industrial Internet of Things Technologies. IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, USA, 1–2.

[99] Sjarov, M., Lechler, T., Fuchs, J., Brossog, M., Selmaier, A., Faltus, F., Donhauser, T., Franke, J., 2020. The Digital Twin Concept in Industry – A Review and Systematization. 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 1789–1796.

[100] Karakra, A., Fontanili, F., Lamine, E., Lamothe, J., 2019. HospiT'Win: A Predictive Simulation-Based Digital Twin for Patients Pathways in Hospital. IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Chicago, USA, 1–4.

[101] David, J.S., 2018. Development of a Digital Twin of a Flexible Manufacturing System for Assisted Learning. Master's Thesis, Tampere University of Technology, Tampere, Finland.

[102] Guo, J., Zhao, N., Sun, L., Zhang, S., 2019. Modular based flexible digital twin for factory design. Journal of Ambient Intelligence and Humanized Computing 10 (3), 1189–1200.

[103] Kunath, M., Winkler, H., 2018. Integrating the Digital Twin of the manufacturing system into a decision support system for improving the order management process. Procedia CIRP 72, 225–231.

[104] Lohtander, M., Ahonen, N., Lanz, M., Ratava, J., Kaakkunen, J., 2018. Micro Manufacturing Unit and the Corresponding 3D-Model for the Digital Twin. Procedia Manufacturing 25, 55–61.

[105] Söderberg, R., Wärmefjord, K., Carlson, J.S., Lindkvist, L., 2017. Toward a Digital Twin for real-time geometry assurance in individualized production. CIRP Annals 66 (1), 137–140.

[106] Chhetri, S.R., Faezi, S., Canedo, A., Faruque, M.A.A., 2019. QUILT: quality inference from living digital twins in IoT-enabled manufacturing systems. Proceedings of the International Conference on Internet of Things Design and Implementation, Montreal Quebec Canada, 237–248.

[107] Havard, V., Jeanne, B., Lacomblez, M., Baudry, D., 2019. Digital twin and virtual reality: a co-simulation environment for design and assessment of industrial workstations. Production & Manufacturing Research 7 (1), 472–489.

[108] Bartsch, K., Pettke, A., Hübert, A., Lakämper, J., Lange, F., 2021. On the digital twin application and the role of artificial intelligence in additive manufacturing: a systematic review. Journal of Physics: Materials 4 (3).

[109] Ko, H., Witherell, P., Ndiaye, N.Y., Lu, Y., 2019. Machine Learning based Continuous Knowledge Engineering for Additive Manufacturing. IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, BC, Canada, 648–654.

[110] Liu, C., Le Roux, L., Körner, C., Tabaste, O., Lacan, F., Bigot, S., 2020. Digital Twin-enabled Collaborative Data Management for Metal Additive Manufacturing Systems. Journal of Manufacturing System.

[111] Mukherjee, T., DebRoy, T., 2019. A digital twin for rapid qualification of 3D printed metallic components. Applied Materials Today 14, 59–65.

[112] Wang, Y., Lin, Y., Zhong, R.Y., Xu, X., 2019. IoT-enabled cloud-based additive manufacturing platform to support rapid product development. International Journal of Production Research 57 (12), 3975–3991.

[113] Joppen, R., Kühn, A., Hupach, D., Dumitrescu, R., 2019. Collecting data in the assessment of investments within production. Procedia CIRP 79, 466–471.

[114] Westermann, G., Finger, S., Giereth, S., Hoffmann, S., Kähler, M., Kölle, V., Popall, M., Reimers, D., Richter, J., Rückriem, K., Schulz, I., Sicorello, S., Thurisch, H., Wendt, S., 2021. Kosten-Nutzen-Analyse: Einführung und Fallstudien. Erich Schmidt Verlag, Berlin.

[115] vom Brocke, J., Hevner, A., Maedche, A., 2020. Design Science Research. Cases. Springer International Publishing, Cham.

[116] Wieringa, R.J., 2014. Design Science Methodology for Information Systems and Software Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg.

[117] Maier, J.F., Eckert, C.M., John Clarkson, P., 2017. Model granularity in engineering design – concepts and framework. Design Science 3.

Biography

Christian Kober, M.Sc. (*1995) studied Industrial Engineering at the University of Applied Sciences Wedel. He joined the Helmut Schmidt University Hamburg in 2021 as research associate and doctoral candidate. His research focuses on economic aspects of Digital Twins in manufacturing. He previously gained experience in the aerospace and automotive industry.

Vincent Adomat, M.Sc. (*1994) studied Industrial Engineering and Business Economics at Leuphana University Lüneburg and International Management and Engineering at Hamburg University of Technology. For pursuing his PhD in Engineering, he joined the Institute of Production Engineering at Helmut Schmidt University in 2020.

Maryam Ahanpanjeh, M.Sc. (*1989) has been a research associate and doctoral candidate at the Helmut Schmidt University since 2021. She received the Master degree in Mechanical Engineering from the Tarbiat Modares University in Tehran. Her main research interests are digitalization of lightweight production and robotics.

Marc Fette, M.Sc., MBA (*1985) has been leading the research group Additive Manufacturing & Lightweight Technologies at chair of Production Engineering in the Department of Mechanical Engineering at Helmut Schmidt University in Hamburg since 2016. He is CEO of the Composite Technology Center of Airbus in Stade and chairman of the Technical Division Aerospace Technologies at VDI.

Univ.-Prof. Dr.-Ing. Jens Wulfsberg (*1959) is head of chair of Production Engineering in the Department of Mechanical Engineering at Helmut Schmidt University in Hamburg since 2001. After having been dean of the Faculty of Mechanical Engineering, he has been the president of the German Scientific Society for Production Engineering (WGP) since 2022.

3rd Conference on Production Systems and Logistics

Development Of A Pre-Competitive Business Model For AI-Based Autonomous Technology Scouting

Stefan Leachu[1], Florian Clemens[1], Volker Stich[1]
[1] *FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany*

Abstract

Technology management can significantly influence the strategic decisions of a company and thus cause success or failure. Basic templates for technology management are technology radars as well as the determination of the technology readiness level (TRL) to be able to evaluate the maturity of newly deployed technologies (e.g., newcomer vs. established). The radars, as well as the TRL, are identified in time-consuming, manual research by subject matter experts from external consultancies. This process is often repeated due to the further development and new development of technologies so that the necessary research becomes an ongoing task. The *TechRad* research project, therefore, aims to automate the identification of the TRL as well as technology radars using web crawling and Natural Language Processing (NLP). To commercialize the pre-competitive prototype, the development of a pre-competitive business model is the goal of this paper. Based on customer analyses, a target group definition is created. Based on user interviews, the precompetitive business model will be detailed in a four-step approach using a business model canvas and a value proposition canvas.

Keywords

Technology Management; Technology Scouting; Technology Radar; Business Model; Business Model Canvas; Value Proposition Canvas

1. Introduction

1.1 Challenge

The number of devices using different digital technologies is increasing and will have more than tripled compared to today by 2025 [1]. Next to that, the number of new technologies is constantly growing [2]. Furthermore, the time until a technology is known to many users is decreasing [3]. It hints that the frequency at which both users and companies are exposed to new technologies is rising. Hence, managing technologies and innovations is a crucial component for entrepreneurial success as it will ensure a company's market position [4]. Staying ahead of the market and managing the sheer number of technologies available is becoming more and more of a challenge for both, large and small enterprises. Being unable to oversee the growing technology market endangers companies to lose their market position and may even result in bankruptcy. Many popular examples such as Nokia and IBM have unveiled the gravity of identifying the right technology trends [5].

1.2 Solution

The *TechRad* research project, therefore, aims to automate the identification of the technology readiness level (TRL) as well as technology radars using web crawling and Natural Language Processing (NLP). The

DOI: https://doi.org/10.15488/12163

solution will improve the technology-scouting process for enterprises, especially SMEs, and assist in building sustainable business growth. To commercialize the prototype developed in the research project, this paper will develop and discuss a pre-competitive open-source business model. The goal of the paper is to enable the commercialization of the prototype either by the project partners or foreign entrepreneurs.

1.3 Structure of the Paper

In the beginning, the paper gives a summary of the implemented prototype, business model, and value proposition canvas. Afterward the general approach for the derivation of the business model is presented in section 3. In section 4 the authors describe the business model and value proposition to commercialize the solution. In the end, there will be a discussion of the feasibility of the commercialization of the solution.

2. Related Work

The proposed business model relies on a solution that enables automated technology monitoring and management. The following paragraph will introduce the solution topics to develop a common understanding of the elements used to develop the business model.

2.1 Autonomous Technology Radar

The solution comprises the gathering, storage, analysis, and visualization of unstructured text data (see Figure 1).

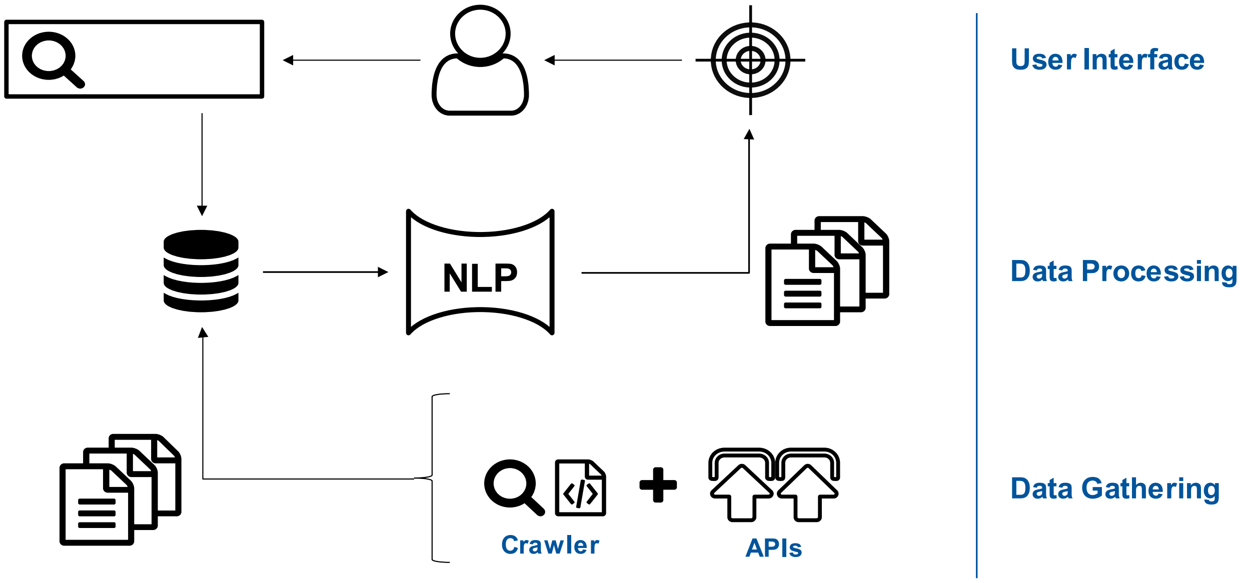


Figure 1: Data Flow of the Solution [6]

The data for the technology radar and TRL is gathered in both ways through API queries and Web-Crawling. Both steps are triggered through search keywords, which are entered by the user. Much information about technologies is available through restricted databases (e.g., *Wikipedia*) which are accessed via API queries. The topic that the user wants to research is passed as an argument to the interfaces to retrieve tailored results. Thus, the query process resembles an up-scaled, automatic version of using the traditional search function on a web page. Next to the API access, an automatic crawler identifies documents from a free web search. Here it needs to be assured to access only content in compliance with current laws and authority. Two distinct kinds of information streams are extracted from the sources. The metadata of the documents as well as information needed to build a crawl index is stored in a permanent database. The full-text versions of the documents are managed differently: Due to considerations of storage space and copyright guidelines, permanent storage of full-text copies is suboptimal, as it may conflict with copyright laws. After the pre-processing step of data analysis, the full copies are discarded, as the necessary information is stored in a high-dimensional vectorized form. It is also necessary to reduce the number of acquired documents as soon as possible. Spam as well as documents with low credibility are discarded. The vectorized texts undergo an analysis step that scores their relevance to the subject and the maturity of the described technology itself. The maturity assessment is performed with NLP. The previously trained model calculates the sentiment of the data based on the spatial distance to other vectorized text samples. That is, if a document in question has similar features to a certain subset of training samples, the resulting spatial distance to these samples will be comparatively low and the model will give it a similar maturity score. The features needed for said assessment are identified by the machine learning algorithm during the training phase. In the end, the results are sent back into the permanent database to aid future related queries. After defining the maturity of the technologies and clustering them into groups in a last step, all the necessary information for building the diagram of the technology radar is available. Supplementing information to the graphic, e.g., hyperlinks to the sources and keywords for further research, finalize the result presented to the user. A second database stores the technology radar in a pool of historical searches to accelerate the fulfillment of future requests. [6,7]

2.2 Business Model Canvas

A business model is a highly simplified and aggregated representation of the relevant activities of a company. It explains how the value creation component of a company generates information, products, and/or services. In addition to the architecture of value creation, the strategic, as well as the customer and market components, are considered to realize the overall goal of generating a competitive advantage. [8]

The business model canvas (BMC) consists of nine elements and is used to define and document a business model [9]:

- **Customer Segments:** The definition of market segments and target groups is the core of any business model, on which all other elements depend.
- **Value Proposition:** The product and its value for the customer must be defined for the respective target group.
- **Channels:** Communication channels and distribution channels are a prerequisite for customers to learn about and purchase the products.
- **Customer Relationships:** The customer relationship describes how the business relationships with the individual customer groups are structured.
- **Revenue Streams:** The goal of every business activity is to generate profit. The revenue streams define how revenues are generated. Together with the cost structure, profitability calculations are possible.
- **Key Resources:** Key resources describe which resources are required to fulfill the value proposition and to serve the customers.
- **Key Activities:** Key activities describe the main activities and competencies to realize the business model.
- **Key Partnerships:** In most cases, a company needs partners to successfully implement a business model. This section lists, for example, the suppliers needed.
- **Cost Structure:** All the costs incurred to realize the business model are compiled and estimated here.

2.3 Value Proposition Canvas

The value proposition canvas is divided into the areas of *customer needs* and *value proposition*. The areas are juxtaposed with each other to show in detail how the value proposition addresses customer needs. A value proposition canvas (VPC) is developed for each customer segment that was identified in the BMC. Customer needs are captured through three perspectives [10]:

- **Customer jobs:** Customer jobs describe tasks and problems that customers want to perform or solve. For this purpose, the functional, emotional, and social needs of the customers are considered.
- **Pain points:** Pain points are challenges that prevent customers from performing customer jobs or solving their problems.
- **Gains:** Gains provide information about how the customer feels, how they describe their sense of achievement, and what they gain from completing the task. Gains also consider functional, emotional, and social aspects.

After the needs and expectations of the customer group have been described, the value proposition is developed against the pains and gains of the customers. Three perspectives are also elaborated for this purpose [10]:

- **Products and Services**: Products and services are selected functions or performance features that support customers in performing tasks and achieving their goals. A distinction is also made between functional, social, and emotional aspects.
- **Pain relievers**: Pain relievers are the aspects of a product that address and resolve the customer's frustrations and problems.
- **Gain Creators**: Gain creators are extras that generate additional and unexpected value and enthusiasm among customers.

3. Methodology and research goal

The overall goal of the underlying research project is to build a prototype to automate the development of technology radars. A business model is needed to commercialize the resulting prototype as a productive application. Therefore, the focus is set on the development of an open-source business model for the automated technology radar. Thus, the applied methodology (see Figure 2) focuses on creating a business model utilizing the BMC and VPC based on the results of qualitative data analysis [11]. The required data was gathered through expert interviews with potential user groups and the project team. Within the next chapter, the derivation of the elements of the business model based on the described approach is explained in detail. Afterward, the gap between the business model and prototype is discussed.

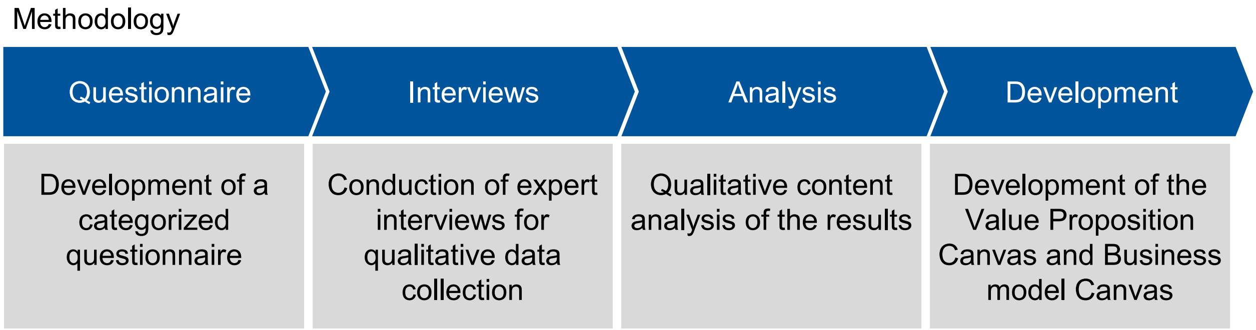


Figure 2: Methodology

4. Pre-competitive Business Model

The BMC and the VPC serve as the basis for the precompetitive business model. As a base of information customer workshops were performed to obtain the following information. In the context of the *TechRad* project, these models are detailed, and the results are described.

4.1 Business Model Canvas

In the following, the nine elements of the BMC are presented (see Figure 3):

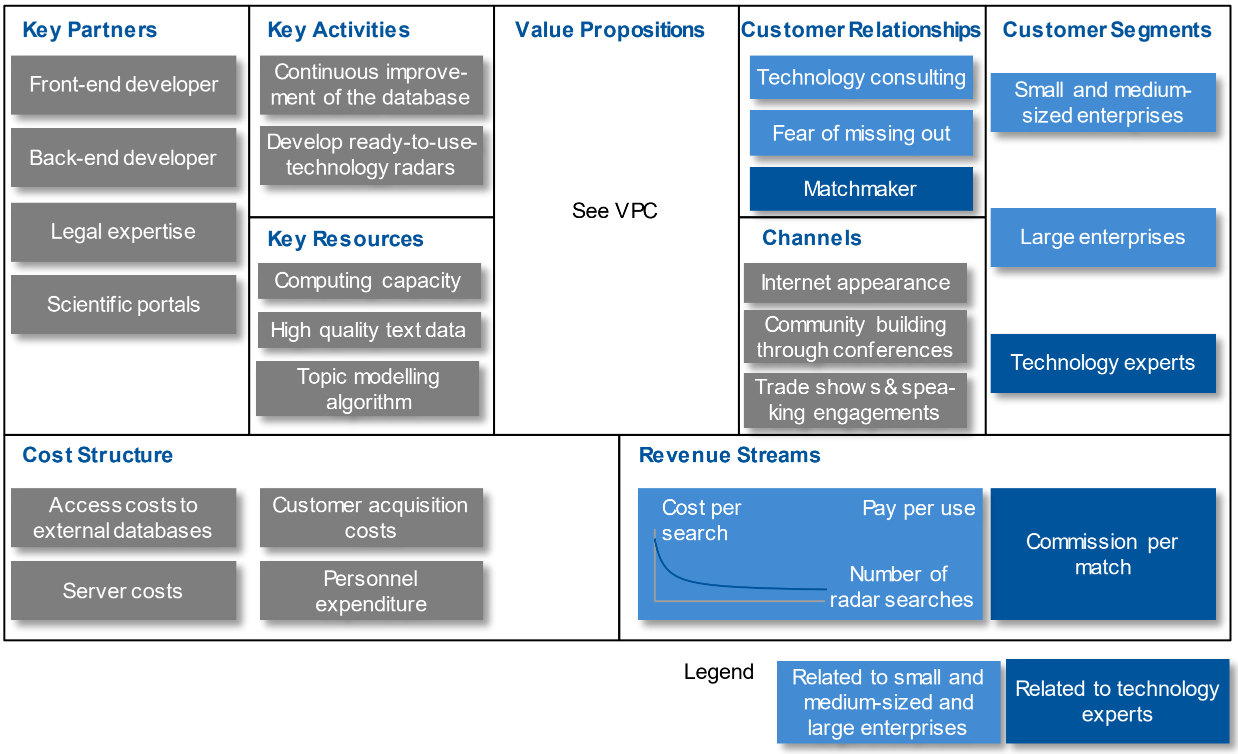


Figure 3: Pre-competitive business model canvas for autonomous technology radars

- **Customer Segments:** The business is targeted to the following groups of customers. They are described by the frequency of use i.e., the number of searches per year. In general, the willingness of SMEs to pay is low due to their infrequent use, i.e., they rarely perform the activity of s*couting* from technology management. Whereas large companies with their technology unit show a high willingness to pay, as these actively perform m*onitoring, scouting,* and *scanning* (cf. Customer jobs in VPC 4.2). Therefore, the potential for saving time is much higher in larger companies. Additionally, technology experts and providers are also addressed with the *TechRad* platform since they can link their expertise on specific technologies to the technology radar.
- **Value Proposition:** The gained value is discussed in detail in the VPC (see Figure 4).
- **Customer Relationships:** The interaction with potential customers is planned in the platform concept with community possibilities. This means that e.g., a forum is provided as an opportunity for interaction within the community. Moreover, the generated technology radars will have a contact person from the platform team for further problems and questions. After passing through the intended use case of the platform, the automated generation of a technology radar, the feedback of users is played back for internal optimization. In general, there is a second customer group, technology experts, and providers. The interaction with this group is done by additional intermediary contracts to act as a broker when customers searching for technologies and get at the same time recommendations for these experts and providers

- **Key activities:** The main aspect of the *TechRad* platform is the distribution of the actual product, the automated Technology Radar. However, the required activities also include continuous development to increase customer satisfaction. On the one hand, the data basis must be constantly expanded, and on the other hand, the algorithms must be maintained and improved. The basis for this can be customer feedback and performance indicators of the algorithms. In addition, a continuous review of the legal principles is required, as external information from a wide range of portals is processed.
- **Key partners:** From an internal perspective, key partners are for front-end development (user interface), for back-end development (AI algorithms), for technological background knowledge, and for legal expertise. External support is provided by portals where technologies and trends are discussed. These range from scientific perspectives (e.g., *ScienceDirect*, *ResearchGate*) to popular scientific publications such as blogs (e.g., *TechCrunch*, *Gartner*) to standards and patents (e.g., *DIN*/ *ISO*).
- **Key resources:** Essential resources can be divided into four groups. First, computing capacity is required in the project to be able to establish constant topicality. In addition, algorithms are also needed for the realization of the project. These include two major AI components, topic modeling algorithm (identification of a topic of a document) and technology readiness level algorithm (identification of a TRL). Basis of all computations is the documents respectively the data basis, which make up the third group. Furthermore, the AI competencies, as well as research competencies are relevant for the design of the platform concept.
- **Revenue Stream:** In parallel to the customer segments, the level of sales is scaled to the size of the company. High revenue is generated by larger companies with frequent use, whereas lower revue is generated by less frequent use by SMEs. The profit is generated by new queries for autonomous technology radars or the update of existing radars. A pay-per-use [12] model is used for this purpose. However, sporadic use is more expensive than frequent use of the platform. The more often a company requests a new radar or updates existing radars, the more favorable these requests will be. In addition, a new branch is opening through the planned referral commissions to technology experts and providers that are associated with the radar.
- **Cost Structure:** Ongoing costs are fundamentally made up of the costs for the data. This includes partnerships with portals and the costs of API accesses. Secondly, the costs of maintaining the infrastructure i.e., server and personnel costs. In addition, there is a further budget for the acquisition of new customers needed.
- **Channels:** Key partners, as well as customers, are reached via several channels. On the one hand, is the online way. A large target group is addressed via blog articles as well as social and display advertisements. For support affiliate programs are used. In addition, search engine marketing and search engine optimization are used to access a larger audience. As a hybrid way of online and offline, community education via conferences is aimed as well as speaking engagements at trade shows. Moreover, direct sales are performed by using demos and training.

4.2 Value Proposition Canvas

In this section, the VPCs six elements are given. Thereby, the elements are split in the customer profile and value map. The VPC shown (see Figure 4) is modeled exemplarily for the customer segment with the highest expected revenue. In this case, these are large enterprises with frequent use of the platform.

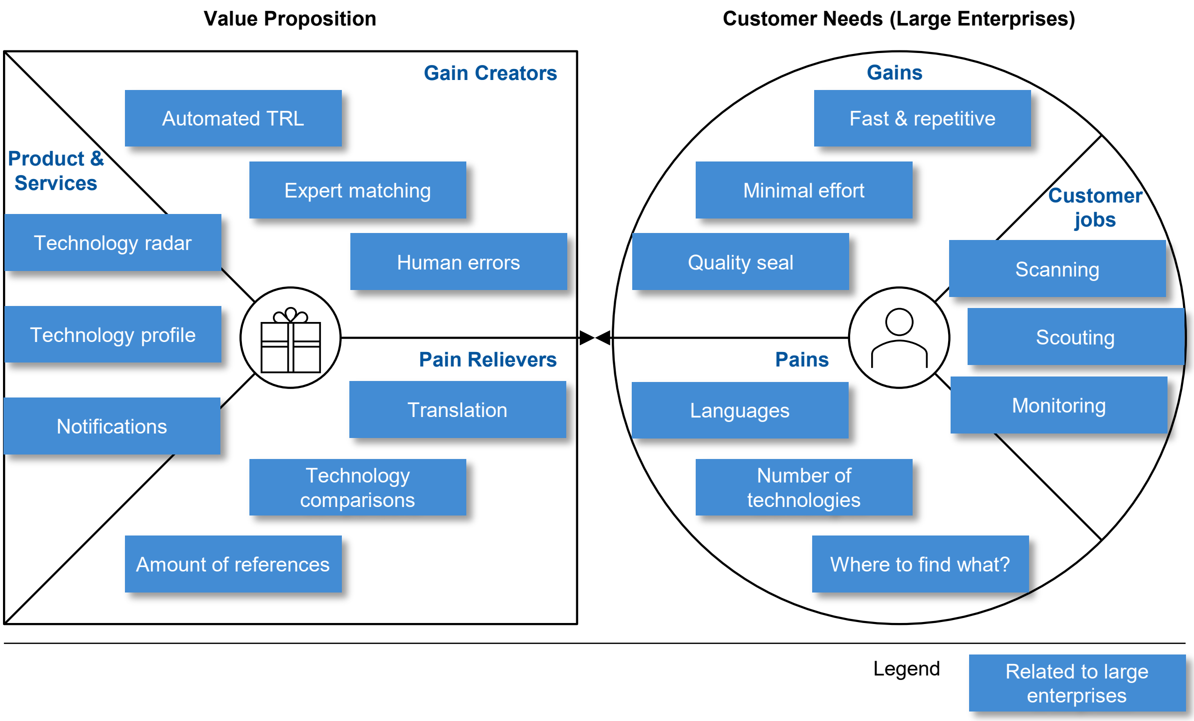


Figure 4: Value Proposition Canvas for the customer segment of large enterprises

Customer Needs:

- **Customer jobs:** Customers generally want to validate their trend analyses in the context of technologies. There are three main jobs that must get done that are part of the technology management processes. First, there is *technology scanning*. This job aims to discover new technology trends. The target image is broadly diversified where the whole market is scanned to early identify new technologies. Thereby, in the business context competing technologies from the perspective of a company are identifiable. Second, there is *technology monitoring*. This job covers monitoring specific technologies to track their evolution. It aims at the one hand to determine the right time to use a technology, on the other hand it is used to observe concrete technologies from competitors. Third, there is *technology scouting* which is used to point out a suitable technology for a specific use case. In other words, it performs a detailed search for predefined criteria. On this base, it can be checked whether technology is appropriated for a purpose. Besides, all gathered results should be manageable and individualizable from a single source.
- **Pains:** When performing the given jobs manually, customers are facing barriers. On the one hand, they reach the limits of their language capabilities. These are comprehension problems caused by foreign languages paired with specialist domain language, for example in patents. Moreover, technology management is recurring and time-consuming work. If you don't know where and how to look up and start, it costs even more time. However, missing technology expertise, in general, leads to the problem of not knowing whether one's status quo in the company also corresponds to the state of the art. It is therefore unclear which technology must be used to hold one's own against competitors on the market. This is how the experience gap arises when a solution is already in use, but the company is not satisfied with it and does not know what alternatives are available.
- **Gains:** The aspects of quick evaluations of the results, which are provided recurrently and continuously, predominate the gains since technology management in general needs a lot of resources (personnel as well as experience). Therefore, a major gain is to facilitate easier access to these results to enable a simple and fast overview of technologies. Moreover, these intense resources cause human errors which must be reduced because every TRL assessment is manually performed

and leads to a summation of errors. This implies the need for a quality seal when results are gathered to ensure trust in the results.

Value Proposition:

- **Products & Services:** The focus of the platform is on the identification and comparison of technologies. The automated generation of technology radars is mainly used for this purpose. The automated identification of TRLs is also fundamental to be able to generate individual radars with any technology. For each technology, profiles are generated in any language to make foreign languages and domain language understandable. In addition, the user experience is enhanced with additional functions. Personal comments are possible for each individually generated radar. The generated radars can be saved and opened again. Also, suggestions are given to already generated radars from the community, which could fit personal preferences. Moreover, special references can be white and blacklisted to customize the automated rating processes. Technologies can also be compared in pairs. In addition, subscription or newsletter functions are available to send emails or customized notifications in the event of changes in trends or technology. Next to the comparison of technologies, the product provides the possibility to display technology experts in the radar and supports the establishment of contact with them.
- **Pain Relievers:** The following problems are addressed by the *TechRad* platform. First, it reduces the time required for technology management, especially research. This is achieved by automatically crawling the internet for sources and interpreting them in the same way to present domain language and foreign languages in a comprehensible way. Furthermore, human errors in the interpretation of these results are reduced, as the identification of a TRL is automated by a standardized process. Overall, quick, and easy access to technology research and its results is offered, making the current state of the art and technology trends transparent.
- **Gain Creators:** The added value of the product is created by automating the previously manual processes, which are time-consuming and require experience. Software packages for technology management already exist to generate technology radars, but the evaluation of the technologies is still the responsibility of the user. Therefore, alternative products are outperformed by *TechRad*, since no continuous and automated TRL determination exists yet and *TechRad* enables this. Similarly, there are no summarized technology profiles, which are generated by *TechRad* in basic language. All information is based on a variety of references, so a wide database exists. Conversely, these sources can also be viewed in the technology radar to enable the traceability of the TRL assessment. In conclusion, information extraction is automated and brought to the users in the broad masses cost-effectively and understandable for everyone without increased effort.

5. Discussion

The business model is based on the assumption that the results of the technology radar are of high quality. That means that they meet the expectations of an expert in the respective domain. SCHUH ET AL. have shown that the results of *TechRads* prototype are acceptable, but the quality still needs improvements [7]. Moreover, not only the quality, but also the time factor proved to be essential in the user workshop. The willingness to pay for the *TechRad* platform has been shown in the fast delivery of the results. The users have specified fast in the range of up to one hour. Due to the fact that the processing in the current prototype phase takes about one day, further development is required in the context of efficiency. However, if the quality and efficiency requirements are not met, commercialization would not become an option.

Furthermore, an autonomous technology radar improves the technology scouting process by reducing the research efforts. This is significant for large companies with technology and innovation management departments. Small companies do not make use of such processes because they do not have technology management departments. Rather, they monitor new technologies until they are mature and then use them to maintain their competitiveness. To ensure that the business model also offers incentives for SMEs, monitoring functions must be developed in further research.

Based on the qualitative content analyses of the results from the expert interviews also experts have been identified as a user group for a technology management platform including the use of autonomous technology radars. The prototype does not include functions to match experts with SMEs or large companies yet, why there is also a need for further development and research in this area.

In summary, data is essential for business models based on AI applications. Therefore, aspects from the key activities, partners, and resources are required as a blueprint in any business models from this area. In fact, these are the continuous further development of the algorithms, legal framework conditions (especially in the European Union), the application infrastructure (frontend and backend), and the IT infrastructure to support the product. Overall, the technical competencies are always needed to implement the goal efficiently and sustainably.

Acknowledgements

This project has been funded by the Leitmarkt.NRW program (EFRE-0801386) and the European Union in the European Regional Development Fund (EFRE) under the name *TechRad*. The authors wish to acknowledge the EFRE for their support. We also wish to acknowledge our gratitude and appreciation to all the *TechRad* project partners for their contribution during the development of various ideas and concepts presented in this paper.

References

[1] Statista, 2022. Number of IoT devices 2015-2025 | Statista. https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/. Accessed 28 January 2022.

[2] Parasuraman, A., 2000. Technology Readiness Index (Tri). Journal of Service Research 2 (4), 307–320.

[3] Statista, 2022. Informationstechnologien - Entwicklung bis 50 Millionen Nutzer | Statista. https://de.statista.com/statistik/daten/studie/298515/umfrage/entwicklung-ausgewaehlter-informationstechnologien-bis-50-millionen-nutzer/. Accessed 28 January 2022.

[4] Klappert, S., Schuh, G., Aghassi, S., 2011. Einleitung und Abgrenzung, in: Schuh, G., Klappert, S. (Eds.), Technologiemanagement. Handbuch Produktion und Management 2. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 5–10.

[5] Gochermann, J., 2020. Technologiemanagement: Technologien Erkennen, Bewerten und Erfolgreich Einsetzen. Springer Fachmedien Wiesbaden GmbH, Wiesbaden, 53 pp.

[6] Schuh, G., Hicking, J., Stroh, M.-F., Benning, J., 2020. Using AI to Facilitate Technology Management – Designing an Automated Technology Radar. Procedia CIRP 93, 419–424.

[7] Schuh, G., Hicking, J., Stroh, M.-F., Benning, J., Gnanaraj, C., 2021. Feasibility Analysis of Entity Recognition as a Means to Create an Autonomous Technology Radar, 9 pp.

[8] Wirtz, B.W., 2021. Business Model Management: Design - Instrumente - Erfolgsfaktoren von Geschäftsmodellen, 5., aktualisierte und erweiterte Auflage ed. Springer Gabler, Wiesbaden, 448 pp.

[9] Osterwalder, A., Pigneur, Y., 2013. Business model generation: A handbook for visionaries, game changers, and challengers. Wiley&Sons, New York, 278 pp.

[10] Osterwalder, A., Pigneur, Y., Bernarda, G., Smith, A., Papadakos, P., 2014. Value proposition design: How to create products and services customers : get started with. Wiley, Hoboken, NJ, 290 pp.

[11] Mayring, P., Fenzl, T., 2019. Qualitative Inhaltsanalyse, in: Baur, N., Blasius, J. (Eds.), Handbuch Methoden der empirischen Sozialforschung, 2. Aufl. ed. Springer VS, Wiesbaden, pp. 633–648.
[12] Gebauer, H., Saul, C.J., Haldimann, M., Gustafsson, A., 2017. Organizational capabilities for pay-per-use services in product-oriented companies. International Journal of Production Economics 192, 157–168.

Biography

Stefan Leachu (*1996) is a researcher at FIR at RWTH Aachen University in the department of Information Management since 2021. With his degree in computer science, he enjoys linking the technical issues of various technologies with the value-added potential that can be achieved in the business world.

Florian Clemens (*1993) received his M. Sc. degree in industrial engineering at University Paderborn. He is a researcher at FIR at RWTH Aachen University since 2019 in the department Information Management, working on the transfer of AI to production.

Prof. Dr.-Ing Volker Stich (*1954) is head of the Institute of Industrial Management (FIR) at the RWTH Aachen University since 1997. Prof. Dr.-Ing Volker Stich worked for 10 years for the St. Gobain-Automotive Group and led the management of European plant logistics. Also, he was responsible for the worldwide coordination of future vehicle development projects.

3rd Conference on Production Systems and Logistics

Software-based Identification Of Adaptation Needs In Global Production Networks

Günther Schuh[1], Andreas Gützlaff[1], Tino X. Schlosser[1], Niklas Rodemann[1], Nicholas Haak[1]
[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, 52074 Aachen, Germany*

Abstract

Internal and external influencing factors force companies to adapt their production networks to changing conditions, which entails a high level of complexity. To be competitive in the future, manufacturing companies have to minimize the required adaptation time between the occurrence of a change and the implementation of an adaptation. While some approaches deal with modelling and evaluating network configuration, there is a lack in identifying the need for adaptation. In practice, the creation of scenarios is often based on the experience and knowledge of the network designer. This paper presents an approach to systematically link perceived key figure changes to possible adaptation alternatives in network configuration. For this purpose, the relevant objects for network adaptations are first defined and adaptation alternatives are systematically described. Subsequently, these are combined with a set of key figures to derive suitable adaptation alternatives depending on their development. The approach is further implemented in a software-based prototype that enables the automated generation of adaptation alternatives in response to perceived changes and provides the user with a listing of possible alternatives prioritized by their utility. The validation with company data demonstrates that by earlier and automated identification of possible configuration adaptations, the adaptation time to changes can be reduced and the generated scenarios are less dependent on the individual experience of the user.

Keywords

Global Production Networks; Network Design; Network Configuration; Adaptation; Optimization Model

1. Introduction

The majority of manufacturing companies of all sizes and industries operate globally in the form of global production networks [1]. These production networks are often historically grown and exposed to a multitude of influencing factors [2]. These internal and external influencing factors are dynamic and require adaptations in the design of the global production network, which is a complex challenge for companies [3]. However, in the face of increasing competitive pressure and dynamics, the ability to adapt to changes is a necessary prerequisite for companies to remain successful in the future [2]. The capability and ability of the footprint to regain a stable state after changes or disruptions is termed network resilience [2]. To improve resilience in global production networks, faster detection of adaptation needs and responsive countermeasures are required [4]. The time required to adapt to a change is divided into three parts and called hysteresis [5]. The first latency period between the occurrence of change until the change is perceived, followed by the latency period until a need for change is identified, and finally the planning latency until the adaptation is implemented [6]. To shorten the adaptation time, the network planner has to be able to react faster in the second part of hysteresis and choose the appropriate adaptation alternative despite the mentioned

DOI: https://doi.org/10.15488/12169

complexity. Therefore this paper focuses on decreasing the time between the perception and identification of a need for change, which can be achieved by the creation of transparency and standardization [7]. The use of company data offers the possibility to record key figures and to present their changes transparently [8,9]. A systematization of adaptation alternatives further supports the selection of a possible response to change [9]. Previous research of the authors systematizes adaptation needs and describes concrete adaptation cases, which now need to be linked to the key figures [10]. In order to provide the network planner with decision support, interactive tools are useful to make the complexity of the planning task manageable [11]. Accordingly, this paper aims at reducing the design complexity of global production networks from a network perspective by combining identified changes and possible network reactions. For this purpose, an indicator-based systematic method is presented to reduce hysteresis by linking the adaptation cases with quantified influencing factors to identify the appropriate response to changes. The approach is further implemented in a software-based prototype that enables the automated generation of adaptation alternatives in response to perceived changes. Prioritization of the adaptation alternatives supports the user in the selection of network adaptations to be considered in more detail.

2. State of the art

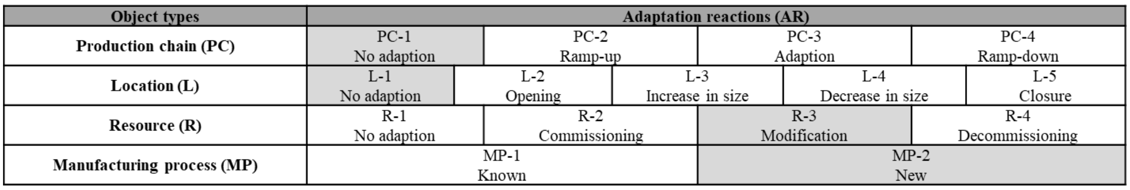
In this context, research approaches regarding structural adaptations of network configuration and adaptation time in network design should be considered in particular. The most current and relevant approaches are presented in the following. WIEZORREK presents an approach for integrating a continuous decision process. Within the framework of permanent monitoring, this process records relevant influencing factors and thus addresses the early identification of the need for adaptation [12]. SCHUH ET AL. provide a reference process for the continuous design of global production networks. The process uses the performance of the production network as a decision basis for identifying the need for adaptation [13]. An approach based on Big data techniques for optimization of network design is presented by GÖLZER ET AL. Within the approach generic planning cases for planning, executing, and validating adjustments are proposed [14]. NEUNER provides a reference framework for the configuration of global production networks considering uncertainty. In the process, uninfluenceable factors are determined and structured according to target variables. These serve as the basis for the evaluation of the configuration alternatives [15]. Some authors use key figures to determine necessary adjustments or to evaluate global production networks. RITTSTIEG examines the factors influencing the performance of production sites. These are quantified by a comprehensive system of key figures [16]. The performance of the production network, as well as environmentally induced adaptation needs, are considered by SAGER. He describes an approach for configuring global production networks by using the concept of selective key figures. Both strategic and operational metrics are used to compare possible adaptation needs in the network configuration [17]. Few authors attempt to handle the complexity of the planning task by implementing interactive software. The solutions developed by SCHUH ET AL. and MOURTZIS ET AL. focus on identifying optimal network configurations based on decisions about the allocation of resources and tasks in the production network, but do not deal in detail with adaptation alternatives to changing influencing factors [18,19]. In summary, approaches to adapting network design as well as the elaboration of key figures can be found in the literature. However, a detailed consideration of the derivation of adaptation needs based on identified changes to shorten the adaptation time is lacking.

3. Conception of the approach

Based on an already existing method for systematizing adaptation cases, chapter 3.1 presents how identified key figure changes can be linked to the adaptation cases. Subsequently, chapter 3.2 prepares the integration into a software tool by creating a data model and introducing the object of the strategic unit. Finally, in chapter 3.3 an optimization model is presented to prioritize the adaptation alternatives.

3.1 Systematized derivation of adaptation alternatives

Within preliminary work, the authors developed an approach to systematize adaptation cases for the design of global production networks. The approach describes each possible design case in the network configuration and allows to structure decisions for a generic production network. A production network is represented by the superposition of several node-edge models, each representing the subnetworks of the product families. Accordingly, the edge of a subnetwork can be understood as the flow of a product between locations in the production network. This flow is referred to as the production chain and can be changed specifically in an adaptation reaction. In addition to the production chain, other network objects are modelled that are relevant for the adaptation. These are the locations of the company with their resources and the manufacturing processes. The adaptation alternatives of the entire production chain result from combinations of the adaptation possibilities for the described object types of the production network. For this purpose, the individual adaptation options for each object type are first defined and bundled in a configuration framework. For example, the production chain has four adaptation options *No Change*, *Ramp-Up*, *Adaptation* and *Ramp-Down*. By linking the adaptation options of each object type, 160 combinations of potential adaptations are obtained for the production chain. Due to internal dependencies and contradictions, these are further reduced to 61. Each of these adaptation cases can be identified by a code resulting from a concatenation of the individual codes. For example, the code 1132 means the modification of a resource and emergence of a new production process without the change of production chain or location (see Figure 1). [10]



<table>
<tr><th>Object types</th><th colspan="20">Adaptation reactions (AR)</th></tr>
<tr><td>Production chain (PC)</td><td colspan="5">PC-1
No adaption</td><td colspan="5">PC-2
Ramp-up</td><td colspan="5">PC-3
Adaption</td><td colspan="5">PC-4
Ramp-down</td></tr>
<tr><td>Location (L)</td><td colspan="4">L-1
No adaption</td><td colspan="4">L-2
Opening</td><td colspan="4">L-3
Increase in size</td><td colspan="4">L-4
Decrease in size</td><td colspan="4">L-5
Closure</td></tr>
<tr><td>Resource (R)</td><td colspan="5">R-1
No adaption</td><td colspan="5">R-2
Commissioning</td><td colspan="5">R-3
Modification</td><td colspan="5">R-4
Decommissioning</td></tr>
<tr><td>Manufacturing process (MP)</td><td colspan="10">MP-1
Known</td><td colspan="10">MP-2
New</td></tr>
</table>

Figure 1: Adaptation reactions in the configuration of global production networks [10]

In order to identify the appropriate adaptation reaction to internal and external influencing factors, the systematized adaptation alternatives described above have to be linked to change drivers. RITTSTIEG and other authors have developed extensive collections of relevant key figures. For the method and the implemented prototype, 15 key figures were selected that were considered to be generally relevant. However, the method works equally with other key figures, which should be selected on a company-specific basis. The linking of the selected key figures is done by analyzing for each adaptation case to what extent it is suitable to counteract deterioration of the key figures. The evaluation is carried out by company experts. Figure 2 shows the section of a general example and represents the interrelations as a table. For adaptation case 1121 it is deduced that it could potentially be used to counteract deteriorations in capacity utilization, area utilization, volume flexibility, or route flexibility. This potential is determined for all adaptation cases. Thus, starting from the deterioration of a key figure, all potentially suitable adaptation cases are captured.

	PC	L	R	MP	Quality rate	Transport time *internal*	Capacity utilization *bottleneck*	Capacity utilization *overcap.*	Area utilization *bottleneck*	Area utilization *overcap.*	Employee product-ivity	Production costs	Material costs	Volume flexibility	Machine flexibility	Transport volume flexibility	Route flexibility	Transport time *supplier*	Transport time *market*
1	1	1	1	2	1	0	1	0	0	0	1	1	0	1	1	0	0	0	0
2	1	1	2	1	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0
3	1	1	2	2	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
·	·	·	·	·															
61	4	5	4	1	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0

PC: Production chain **L**: Location **R**: Resource **MP**: Manufacturing process (1) interrelation (0) no interrelation

Figure 2: Linking the key figures with the adaptation cases (example)

Combined, the individual adaptation cases result in an adaptation alternative for the entire production network. However, this approach still has application-related gaps that need to be closed. In a production network, there are numerous network objects that are all interdependent. Key figure changes can occur simultaneously in several objects. Therefore, in the following a consideration of a multitude of key figures is enabled and adaptation alternatives are generated in an object related way. In addition, interdependencies between production chains have to be taken into account by determining the adaptation reactions of the entire production network simultaneously instead of considering the individual production chains successively.

3.2 Division of the production network into strategic units

In order to provide the network planner with software-based decision support, a data model has to be set up that contains the object types for defining adaptation cases. In addition, the model of the production network is extended by the suppliers and the sales market to take external key figures into account. Further, transport routes are integrated into the network that link two locations with each other as well as suppliers and sales markets with a location. The resulting data model is implemented as a class diagram based on the Unified Modeling Language (UML) and visualized in Figure 3.

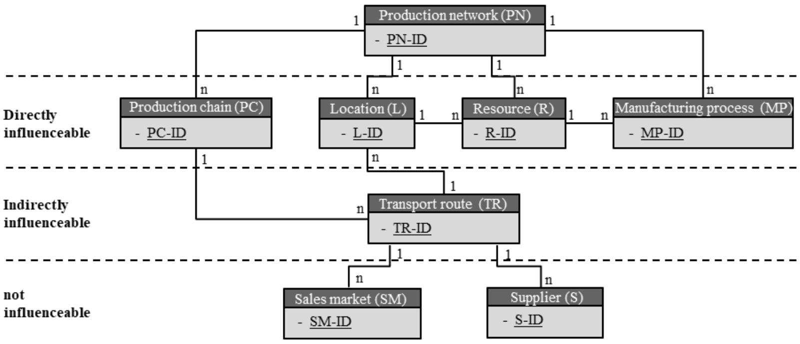


Figure 3: UML data model for configuring global production networks

For an assignment of design measures in the production network, a distinction is made in the data model regarding the influenceability of the network objects. While the influenceable objects also serve as the output of an adaptation reaction, the non-influenceable objects are only seen as input, i.e. they can only be used to identify a need for action. The network objects that are addressed by the adaptation cases, i.e. the production chains, locations, resources and manufacturing processes, are classified as directly influenceable. On the other hand, the sales market and the suppliers cannot be influenced. These are defined as exogenous in the model. In between are the transport routes, which connect the locations internally with each other as well as locations with suppliers or customers. Since sites and production can be influenced, the transport routes can also be addressed indirectly through design measures. However, due to the dependency on the exogenous objects of the network, their adaptation possibilities are limited. The selection and distribution of suitable adaptation measures in the production network is based on the reference process for the continuous design of production networks according to SCHUH ET AL. [13]. In the reference process, a network configuration for the entire production network is determined on a tactical level by decomposing the network into the individual value streams. For these value streams, possible scenarios are developed and evaluated, checked for dependencies, and finally selected. In this work, the production network is grouped into strategic units, each containing a value stream and the network objects relevant to the value stream. The strategic units contain all strategic decisions of relevant objects, which are used for identification as well as for the implementation of an adaptation. By decomposing the production network into such units, several of the generic adaptation cases can be assigned to the production network at the same time, in that each strategic unit receives exactly one adaptation case if action is required within the strategic unit. Thus, the adaptation of the network is no longer dependent on the successive consideration of individual production chains, but all production chains can be considered simultaneously. In addition, several network objects of the same class can be addressed within a production chain. For example, adaptation alternatives can be identified that react simultaneously to key figure changes from two different locations and select suitable adaptation cases in each case. The structure of a strategic unit is shown in Figure 4.

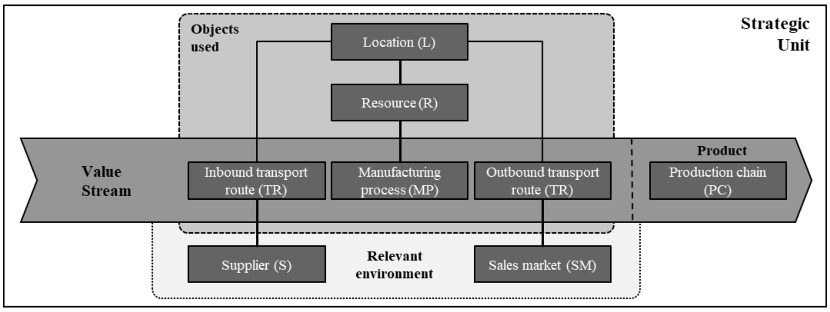


Figure 4: Model of a strategic unit

At the center of each strategic unit is the value stream, which always relates to a specific product and is therefore part of a production chain. The value stream begins with the inbound transport route, which delivers the product to the respective location. This transport route starts either at a supplier or at a company-internal location. At the core of the value stream is the manufacturing process used. As a value-adding element, the manufacturing process requires a suitable resource for its execution, which is located at a site. The output of the value stream is the outbound transport route, which leads either to another location (and thus to another strategic unit) or to the sales market. Each of the network objects mentioned occurs exactly once in a strategic unit, so that there are at least as many strategic units as there are manufacturing processes in the company. If a manufacturing process is linked to several transport routes at the input or output, the number of strategic units is even higher. Each strategic unit can be assigned to one of the 61 generic adaptation cases by addressing the network objects that can be directly influenced. The need for action is determined based on the key figure changes that occur in the network objects of the strategic units. Thus, there is a strategic dependency within each strategic unit, as the objects influence or constrain each other in the choice of an adaptation action. Outside the strategic unit, other primarily structural interdependencies have to be considered. For example, the same network objects may occur in several units, i.e. a site could have two resources, each has two manufacturing processes. In this case, four strategic units would be created, so that the site as well as the resources would occur various times. Therefore, when designing all strategic units, it is necessary to ensure that the actions of the network objects are unique.

3.3 Optimization model

3.3.1 Structure of an optimization model

The method presented in this paper aims to identify adaptation needs at an early stage by automatically generating adaptation alternatives to changes in key figures. A decision problem arises about which adaptation cases are assigned to which strategic units, so that the action requirements are met in the best possible way considering the restrictions in the network. Decision problems can be solved with qualitative, quantitative and combined methods. In contrast to qualitative methods, quantitative ones are based on objectively measurable criteria and serve as a rule for the optimization of a target value by parameter variation [20]. The presented decision problem in this paper considers objectively measurable ratio changes. To enable an automatic identification of adaptation alternatives subjectivity is to be avoided and a quantitative evaluation method is appropriate. For mapping the decision problem a mathematical optimization model is chosen to deal with the optimization of functions under constraints. An optimization model is a formal representation of a decision problem that contains, in its simplest form, at least one set of alternatives and an objective function evaluating them. The model is developed to be able to determine optimal proposed solutions using appropriate procedures. In its basic structure, an optimization model consists of an objective function to be optimized, a variable vector describing the alternative courses of action, and a restriction system consisting of several constraints that restrict the solution space and define the range of values of the decision variables [21]. An important special case of mathematical optimization are linear optimization models. A linear optimization model exists if the variables are not multiplied by each other and no variables are found in exponents. If it is possible to put a mathematical optimization model into this linear form, enormous advantages arise, since fundamental efficient methods for linear models have been developed. This enables to communicate the mapped problems as well as problem instances to standard software known as solvers, which solve the problem instance optimally. These solvers have increased their performance enormously in recent years, allowing them to solve increasingly complex problems more efficiently [22]. For this reason, a linear optimization model is chosen. By adding secondary conditions to the model, the solution space of the problem is narrowed down and restrictions from the structure of the production network considered [22]. In the following, the optimization model is described in detail.

Objective function:

$$max \quad N(x) = \sum_{i\in S}\sum_{j\in P} n_{ij}x_{ij} \tag{1}$$

Secondary conditions:

$$\sum_{j\in P} x_{ij} = 1 \,\forall\, i \in S \tag{2}$$

$$n_{ij} = \sum_{k\in K'} r_k * e_{jk} * z_{ik} \;\; \forall\, i \in S, j \in P \tag{3}$$

$$x_{ij} \in \{0,1\} \;\; \forall\, i \in S, j \in P \tag{4}$$

$$x_{ij} + x_{lm} \;\leq 1 \;\; \forall\, (i,j,l,m) \in K \tag{5}$$

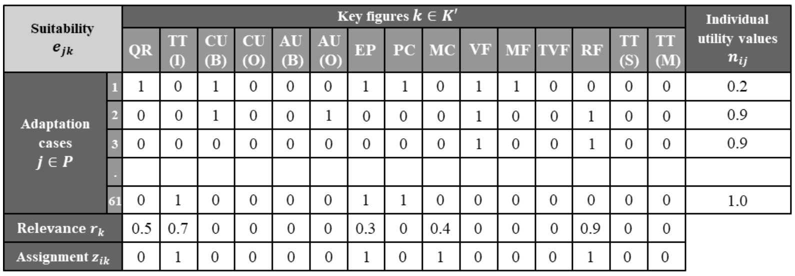
$$x_{ij} = 0 \;\; \forall\, (i,j) \in R \tag{6}$$

Table 1: Overview of the elements of the optimization model

Category	Symbol	Description
Index	*i,l*	Index of a strategic unit
	j,m	Index of a generic adaptation case
	k	Index of a key figure change
Decision variable	*x*	Binary variable indicating the assignment of an adaptation case to a strategic unit
	e	Binary variable for assigning an adaptation case to a key figure change
	z	Binary variable for assigning a key figure change to a strategic unit
Coefficient	*n*	Matrix of the utility values of the adaptation cases for each strategic unit
Objective function value	*N*	Total utility value of the adaptation alternative for the production network
Set	*S*	Set of strategic units in the production network
	P	Set of 61 generic adaptation cases
	K'	Set of all key figures
	K	Conflicts between two assignments resulting from the structure of the network
	R	Restrictions that arise from individual specifications of the user

3.3.2 Definition of the objective function

The objective of the method is to generate the best possible adaptation alternatives for the production network, which is captured in the optimization model by the objective function (1). The first constraint (2) specifies that only one adaptation case $j\in P$ can be assigned to each strategic unit $i\in S$. This condition is needed so that reasonable solutions can be determined. Overall, the adaptation cases are to be assigned to the strategic units in such a way that the total utility value of all assignments is maximized. To achieve this, the first step is to calculate the individual utility values for all combinations of potential assignments. These individual utility values serve as parameter *n* for calculating the total utility value of an adaptation alternative. A utility analysis is used to calculate the individual utility values. In practice, this is a frequently applied procedure for the evaluation of alternative actions [20]. In this process, the alternatives are ordered according to the preferences of the decision maker concerning a multidimensional target system [23]. In the context of the method, the action alternatives are the adaptation cases of the production network. The target system results from the multidimensional key figure system. The utility of an adaptation case is calculated by how many key figure changes (considering their relevance) can be addressed. Formula (3) represents the individual benefits of each adaptation case for each strategic unit. K' is the set of all key figures and r_k is the relevance of the development of a key figure $k \in K'$, which can be calculated via the relative deviation of the key figure development. The binary variable e_{jk} indicates whether an adaptation case $j \in P$ is suitable for addressing a key figure change *k*. This information can be derived from the table in Figure 2. The individual utility values *n* have to be calculated for each strategic unit, since the key figure changes are assigned to the objects of the production network and several strategic units do not exclusively contain the same objects. Therefore, the action required per strategic unit may vary. In order to take this into account, the utility analysis is complemented by the additional binary variable z_{ik}, which indicates whether the key figure change *k* occurs within the strategic unit $i \in S$. For the implementation of the method in a software demonstrator, the strategic unit is included as a network object in a database so that the relationship to the other objects can be retrieved and used to automatically determine z_{ik}. For individual strategic units, the calculation of individual utility values can be performed using a table (see Figure 5).



Suitability e_{jk}		Key figures $k \in K'$															Individual utility values n_{ij}
		QR	TT (I)	CU (B)	CU (O)	AU (B)	AU (O)	EP	PC	MC	VF	MF	TVF	RF	TT (S)	TT (M)	
Adaptation cases $j \in P$	1	1	0	1	0	0	0	1	1	0	1	1	0	0	0	0	0.2
	2	0	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0.9
	3	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0.9
	.																
	61	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	1.0
Relevance r_k		0.5	0.7	0	0	0	0	0.3	0	0.4	0	0	0	0.9	0	0	
Assignment z_{ik}		0	1	0	0	0	0	1	0	1	0	0	0	1	0	0	

Figure 5: Exemplary calculation of the individual utility values for a strategic unit

In this example, the utility value of adaptation case 61 for the considered strategic unit with code number 1 is calculated as followed:

$$n_{1,61} = \sum_{k \in K'} r_k * e_{61,k} * z_{1,k} = (0.7 * 1 * 1) + (0.3 * 1 * 1) = 1.0 \qquad (7)$$

Thus, adaptation case 61 has the highest individual utility value of the adaptation cases considered in the example. The individual utility values are calculated for each strategic unit and serve as coefficients for the objective function (1) of the linear optimization model. The total utility value *N* is the sum of the individual utility values of all selected adaptation cases. The decision variable x_{ij} serves for the assignment of an adaptation case $j \in P$ to a strategic unit $i \in S$. x_{ij} is a binary variable that takes the value one if an assignment takes place and zero if not. This binarity is expressed in the optimization model by constraint (4). The task of the solver is to determine a value for all assignments so that the objective function is maximized. Constraints (5) and (6) represent restrictions of the network and are explained in the following subchapter.

3.3.3 Restrictions from the network configuration

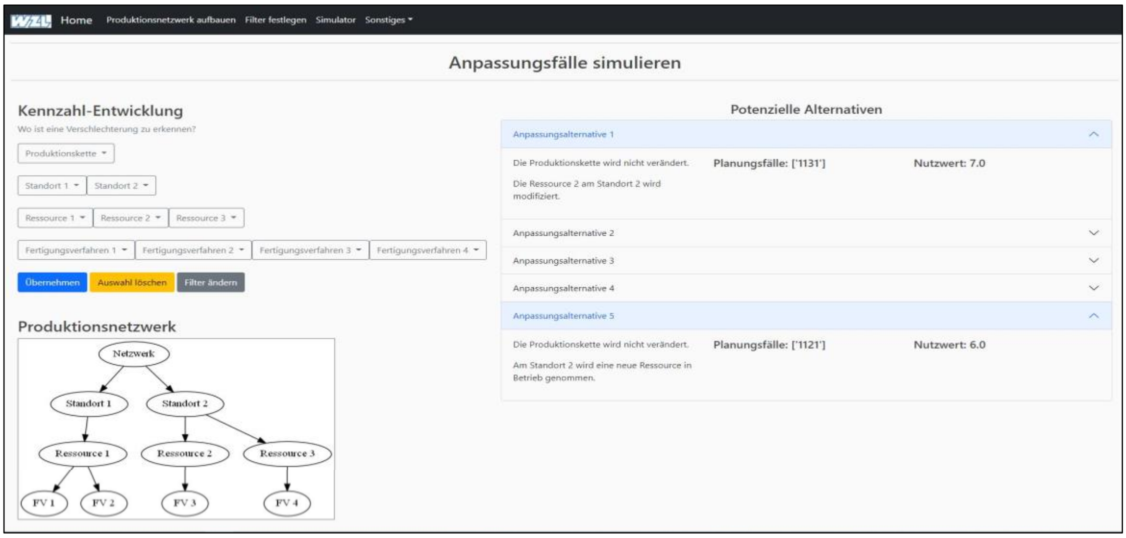
If there are no interdependencies between strategic units, the objective function could easily be optimized by selecting all adaptation cases with maximum individual utility values. In reality, the selected adaptation cases may contradict each other and thus not be feasible. For example, one adaptation case might involve the decommissioning of a resource, while another adaptation case involves only a modification for the same resource. Therefore, structural interdependencies between the strategic units of the production network have to be considered when selecting adaptation cases. These interdependencies are accounted for by constraint (5) in the optimization model. The constraint specifies that two assignments (each between a strategic unit and an adaptation case) cannot both be selected if this constellation is stored in the conflict list *K*. The conflict list *K* is a list of interdependencies between the strategic units of the production network. The conflict list depends on the structure of the production network and the composition of the strategic unit. Therefore it has to be created individually for each production network in advance. In this method, the conflict list is created with the help of an algorithm. All possible assignments are reviewed and checked for dependencies between the strategic units. The basis is the contradictions in the generic actions of the individual network objects. If there is a contradiction, the two actions under consideration cannot be performed within the same object. In preliminary work of the authors, the actions were defined as unique and not overlapping [10]. Therefore, a uniform adaptation reaction has to be selected for a single object. However, there is the exceptional case that an adaptation reaction creates a new network object. Then it is possible to perform any action on the first object as well as to create the new object. For example, a new resource could be put into operation while the old resource is modified. This also applies to the opening of a site and the development of a new manufacturing process. The described constraints are used to consider the structure of the production network. With regard to the application of this methodology, however, it should be possible to generate additional constraints that incorporate strategic guidelines from the company. Thus, guidelines from the network strategy can be considered. For example, location decisions can be dependent on a superordinate strategy, such as the development of a new sales market. In addition, the solution space of the problem can be further restricted by following the guidelines, so that the decision is facilitated. In this optimization model, the user-specific restrictions form the constraints (6), which are not further detailed here.

3.3.4 Iterative solution of the optimization problem and prioritization of the alternatives

After all required coefficients and quantities have been determined, the optimization problem can be set up and solved. Since the optimization model contains a binary decision variable, an integer optimization problem has to be solved. Manual selection of an algorithm is not necessary, since a standard solver (*Coin-or-branch and cut*) is used in the software implementation, which automatically determines a suitable algorithm. Since the adaptation alternatives as results of the method should only serve as decision support for the network planning, the specification of a single adaptation option is not purposeful. Rather, several alternatives for the production network should be generated and listed according to their utility value. This results in a clear solution space of potential alternatives, which are further checked for feasibility and reasonableness. For this purpose, the optimization problem is solved iteratively. In each subsequent iteration the solution of the previous iteration is forbidden. Thus, each iteration provides an additional adaptation alternative for the production network, whose utility is smaller or equal to the utility in the previous iteration. Thus, starting from the second iteration, a new constraint must be included in the optimization model that excludes all previous solutions.

4. Application

The described method was transferred into a software demonstrator, which is structured in the form of a web application and can be operated via any internet browser (see Figure 6). The software demonstrator allows the user to simulate various situations to identify individually tailored adaptation alternatives for a production network. The result is a prioritized list of adaptation options, which the user should then evaluate based on various criteria (effort, cost, risk, etc.) in order to finally adapt the production network. The decision is facilitated by the application, as the solution space is reduced by systematizing the adaptation cases and prioritizing the network alternatives. The software demonstrator was applied and validated at a household appliance manufacturer. The company's network consists of several global locations and has grown historically. There is high potential by adapting to changes such as growing unit numbers in new markets. A key figure system for monitoring the production network consisting of 15 key figures was used for the validation. The key figures were examined for changes using historical data. Five key figures were identified that had changed in different optimization directions. By creating a table to determine the dependencies between the 61 adaptation cases and the key figures, the software demonstrator could be fed with the corresponding data. The resulting list of prioritized adaptation alternatives was validated by a network planner. 5 of the 8 highest prioritized alternatives were classified as realistic adaptation alternatives and could be investigated in a next step with respect to the criteria mentioned above.



Die Produktionskette wird nicht verändert.

Die Ressource 2 am Standort 2 wird modifiziert.

Planungsfälle: ['1131']

Nutzwert: 7.0

Die Produktionskette wird nicht verändert.

Am Standort 2 wird eine neue Ressource in Betrieb genommen.

Planungsfälle: ['1121']

Nutzwert: 6.0

Figure 6: Excerpt of the software tool (anonymized example)

5. Conclusion

The paper presents a methodology and its transfer into a software tool for the identification and prioritization of possible adaptation alternatives in global production networks. This provides decision support to the network planner by narrowing the solution space and suggesting possible adaptation alternatives in an early and automated way. There is currently a need for research in the processing of the data from the systems, which enables the automated calculation of the key figures used. In addition, possibilities for further evaluation of the prioritized adaptation alternatives should be investigated. Factors such as effort, cost, strategic importance, sustainability, and risk of the identified alternatives should be considered to assess the alternatives for feasibility and reasonableness. Further development of the method focuses on reducing the high degree of subjectivity in linking the metrics to the adaptation cases. This could potentially be countered by a feedback learning system, using appropriate machine learning algorithms to adjust the values of the table used to produce more meaningful results.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Ferdows, K., 2018. Keeping up with growing complexity of managing global operations. International Journal of Operations & Production Management 38 (2), 390–402.

[2] Moser, E., Stricker, N., Lanza, G., 2016. Risk Efficient Migration Strategies for Global Production Networks. Procedia CIRP 57, 104–109.

[3] Váncza, J., 2016. Production Networks, in: The International Academy for Production Engineering, Laperrière, L., Reinhart, G. (Eds.), CIRP Encyclopedia of Production Engineering, vol. 147. Springer, pp. 1–8.
[4] Lanza, G., Ferdows, K., Kara, S., Mourtzis, D., Schuh, G., Váncza, J., Wang, L., Wiendahl, H.-P., 2019. Global production networks: Design and operation. CIRP Annals 68 (2), 823–849.
[5] Brokate, M., 1996. Hysteresis and Phase Transitions. Springer New York, New York, NY, 368 pp.
[6] Ferdows, K., 2014. Relating the Firm's Global Production Network to Its Strategy, in: Johansen, J., Farooq, S., Cheng, Y. (Eds.), International Operations Networks, vol. 84. Springer London, London, pp. 1–11.
[7] Schuh, G., Prote, J.-P., Fränken, B., Dany, S., Gützlaff, A., 2018. Reduction of Decision Complexity as an Enabler for Continuous Production Network Design, in: Moon, I., Lee, G.M., Park, J., Kiritsis, D., Cieminski, G. (Eds.), APMS, vol. 535. Springer International Publishing, Cham, pp. 246–253.
[8] Gottmann, J., 2019. Produktionscontrolling. Springer Fachmedien Wiesbaden, Wiesbaden.
[9] Schuh, G., Gützlaff, A., Thomas, K., Rodemann, N., 2020, Framework for the Proactive Identification of Adaptation Needs in the Configuration of Global Production Networks, in: 2020 International Conference on Industrial Engineering and Engineering Management, Singapore, pp. 69–73.
[10] Rodemann, N., Niederau, M., Thomas, K., Gützlaff, A., Schuh, G., 2021. Systematization of Adaptation Needs in the Design of Global Production Networks, in: Behrens, B.-A., Brosius, A., Hintze, W., Ihlenfeldt, S., Wulfsberg, J.P. (Eds.), Production at the leading edge of technology. Springer, pp. 650–659.
[11] Friedli, T., Schuh, G., Lanza, G., Remling, D., Gützlaff, A., Stamer, F., 2019. Next Level Production Networks. ZWF 114 (3), 101–104.
[12] Wiezorrek, A., 2017. Beitrag zur Konfiguration von globalen Wertschöpfungsnetzwerken. Dissertation.
[13] Schuh, G., Prote, J.P., Dany, S., 2017, Reference process for the continuous design of production networks, in: 2017 International Conference on Industrial Engineering and Engineering Management, Singapore, pp. 446–449.
[14] Gölzer, P., Simon, L., Cato, P., Amberg, M., 2015. Designing Global Manufacturing Networks Using Big Data. Procedia CIRP 33, 191–196.
[15] Neuner, C., 2009. Konfiguration internationaler Produktionsnetzwerke unter Unsicherheit. Dissertation.
[16] Rittstieg, M., 2018. Einflussfaktoren der Leistungsfähigkeit von Produktionsstandorten in globalen Produktionsnetzwerken. Dissertation.
[17] Sager, B.M., 2019. Konfiguration globaler Produktionsnetzwerke. Dissertation.
[18] Mourtzis, D., Vlachou, E., Boli, N., Gravias, L., Giannoulis, C., 2016. Manufacturing Networks Design through Smart Decision Making towards Frugal Innovation. Procedia CIRP 50, 354–359.
[19] Schuh, G., Potente, T., Varandani, R., Schmitz, T., 2014. Global Footprint Design based on genetic algorithms – An "Industry 4.0" perspective. CIRP Annals 63 (1), 433–436.
[20] Burggräf, P., Schuh, G., 2021. Fabrikplanung. Springer Berlin Heidelberg.
[21] Domschke, W., Drexl, A., Klein, R., Scholl, A., 2015. Einführung in Operations Research. Springer.
[22] Briskorn, D., 2020. Operations Research. Springer Berlin Heidelberg.
[23] Meyer, H., Reher, H.-J., 2020. Projektmanagement. Springer Fachmedien Wiesbaden.

Biography

Günther Schuh (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University, is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen.

Andreas Gützlaff (*1989) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. He is Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Tino X. Schlosser (*1989) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen and Tsinghua University. He is Team Lead Global Production in the Production Management department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Niklas Rodemann (*1992) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University. Since 2018, he has been a research assistant in the Production Management department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University. His research focuses on the configuration of global production networks.

Nicholas Haak (*1997) studied Mechanical Engineering specializing in Production Engineering at the RWTH Aachen University. He is a Student Research Assistant at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University in the Production Management department.

3rd Conference on Production Systems and Logistics

DISPO 4.0 | Digitalization Of Inventory Calculation In Consumption-Based Material Requirements Planning In The Capital Goods Industry

Alexander Schmid [1,2], Thomas Sobottka [1,2], Wilfried, Sihn [1,2]

[1] *Fraunhofer Austria Research GmbH, Theresianumgasse 7, 1040 Vienna, Austria*
[2] *TU Wien, Institute of Management Science, Theresianumgasse 27, 1040 Vienna, Austria*

Abstract

This paper presents a material requirements planning method that determines optimal safety stock levels using a heuristic optimization, based on a deterministic simulation of stock levels. Material requirements planning is a key competitiveness factor in a volatile, global market environment and is becoming increasingly complex due to the availability of more products, product variants and fluctuating demand. Digitalization offers significant potential benefits for this planning domain, however, tools ready for use in industry applications are still lacking, leading to untapped potential in companies. The approach presented herein investigates available safety stock calculation algorithms, develops a heuristic-based optimization method that determines the best fitting algorithm for each product and optimally parameterizes the algorithm. The method utilizes a deterministic simulation as an evaluation function. A case study for a company in the capital good industry is implemented to evaluate the application potential. The results reflect significantly improved service levels with a minor increase in cost.

Keywords

algorithms; calculation of stock; consumption-based material requirements planning; digitalization; heuristic optimization; inventory calculation; safety stock; safety stock planning; simulation

1. Introduction

Data and information are sometimes referred to as the “oil of the digital age". This increasingly applies to material requirements planning, which is confronted with increasing complexity in a volatile, global market environment and the associated increase in data volumes [1]. Disruptions due to digitalization, smaller batch sizes, fluctuating sales volumes, globalized supply chains and cost pressure are major complexity drivers in material requirements planning [2]. Material requirements planning refers to the coordination of the flow of materials into the company and the stock levels so that the required items are available on time and in the right quality, at the right place [3]. The aim of material requirements planning is to ensure that the company's material supply is economically secure in terms of type, quantity, time and quality [4]. The sub-disciplines of material requirements planning are divided into requirements planning, calculation of stock and purchase order calculation [3], see Figure 1. This paper focuses on the sub-discipline of calculation of stock, specifically on the application of safety stock algorithms in consumption-based material requirements planning. Although many companies view safety stock primarily as a cost driver, safety stocks are the key factor for maintaining a high service level [4]. Optimally defining safety stock levels help on the one hand to increase the service level for the customer, and on the other hand to minimize company-relevant inventory costs [5]. Digitalized, automated planning can achieve significant savings, ensure long-term customer loyalty

DOI: https://doi.org/10.15488/12173

and improve competitiveness [6], and a variety of algorithms exist in consumption-based material requirements planning to improve its efficiency. However, only a very small proportion of mathematical models are applied in day-to-day operations [7].

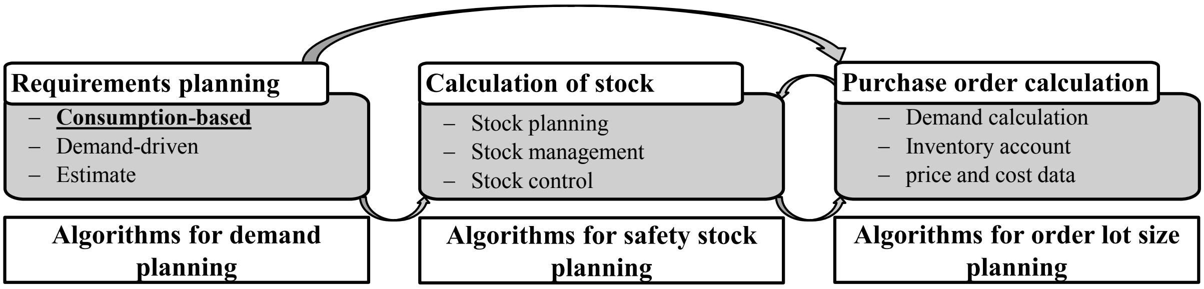


Figure 1: Sub-disciplines of material requirements planning

This paper presents the development of a digital planning tool for material requirements planning and operational purchasing that enables product-specific optimized calculation of stock. The objective is to guarantee the availability of consumption-controlled disposition, considering potential item-specific uncertainties in the supply chain, with the lowest possible safety stocks. For this purpose, a heuristic optimization based on a deterministic simulation is developed as an evaluation function. The potential benefits for optimized safety stock calculation are evaluated in a case study from the capital goods industry. Its relevance for the industry can be justified by the fact that capital goods are increasingly placed at shorter notice and for smaller volumes. Therefore, producers are being demanded short delivery times, a high degree of flexibility and of planning accuracy and this with an increasing variety of articles.

The research hypothesis is that a digital planning method in consumption-based material requirements planning can significantly increase the service level compared to the safety stock determinations practiced today in companies in the capital goods industry. The Design Science Research Methodology according to *Peffers et al.* [8] was applied, supporting both the development of a solution and its communication into application.

The paper is structured as follows: Following the introduction, section 2 provides relevant fundamentals for safety stock planning, while section 3 introduces simulation and optimization for safety stock planning. Section 4 presents the development of the planning method. In the concluding sections 5 and 6, the results are discussed, and an outlook is provided.

2. Background: Safety stock planning

A literature analysis provides an overview of available safety stock calculation algorithms (see Figure 2). Altogether, 16 different methods could be identified. None of the referenced literature considers all algorithms. Figure 3 categorizes the algorithms and outlines their relationships to each other. The algorithms were then characterized and the possible applications in the operational environment of the capital goods industry were evaluated. The procedures marked in dark grey were selected as the most common procedures though a prevalence analysis. Some of these 11 algorithms are already used in Enterprise resource planning (ERP) systems. However, decision-makers in companies lack a basis for deciding which of the safety stock algorithms are most suitable and how to parameterize them optimally and in a product-specific manner.

In real-life settings there are numerous factors that can contribute to uncertainty in material requirements planning [9], such as delivery date deviations, delivery quantity deviations, consumption deviations, supplier quality problems and stock deviations. In order to counteract the occurrence of shortages in materials disposition, safety stocks are used as buffers in materials disposition. Planners have to decide between a high level of service and the associated higher capital commitment and storage costs, and between low stocks and the associated risk of the occurrence of shortages [10].

Algorithms *References*	Uncertain lead time (Theory of Constraints)	Safety stock with dynamic service-levels	Safety stock with target service-level	Safety stock dynamic with target service-level	Calculation according to service-level	Calculation by using the A-B-C/X-Y-Z method	Dynamic safety stock method	Calculation with a service-level of 100%	Calculation by means of the rough estimate method	Calculation by Lagrange method	Calculation by Percent-Fill Method	Calculation by means of double reorder point	Calculation by means of percentage surcharge	Experience-based method	Square Root Law	Calculation by means of portfolio effect
Becker et al., (2014)				x			x									
Brabänder, (2020)					x				x						x	x
Bretzke, (2020)															x	x
Dangelmaier, (2017)						x										
Gudehus, (2006)	x													x		
Gudehus, (2012)	x															
Heiserich et al., (2011)					x				x			x	x			
Lödding, (2016)			x					x								
Luthra, (2011)									x			x	x			
Nyhuis u. Wiendahl, (2012)			x					x								
Pfohl, (2018)									x			x	x			
Radanasu, (2016)					x				x					x		
Schmidt et al., (2012)	x	x			x				x							
Schönsleben, (2020)					x				x							
Schuh & Schmidt, (2014)	x								x							
Stead, (1990)						x										
Thomopoulos, (2016)					x					x	x					
Wannenwetsch, (2014)									x							
Wiendahl, (2020)					x				x							

Figure 2: Literature allocation to safety stock procedures

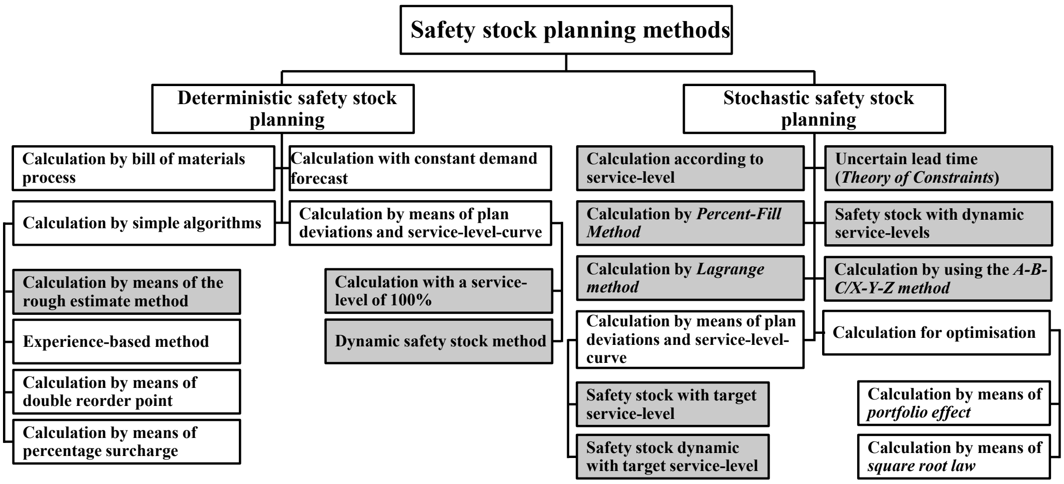


Figure 3: Overview of safety stock algorithms

3. Heuristic optimization, simulation and simulation-based optimization of safety stocks

Optimization is the process of finding the best possible solution for the objective – in this case, maximum service level with minimum inventory – with the help of mathematical operations. The optimization can either be implemented as a mathematical optimization program or as an algorithm that uses an evaluation function to achieve the objective. For complex real systems, some form of simulation is often useful for the evaluation function – in this case, the stock is deterministically simulated over time for the different methods to analyze, how a chosen stock calculation method affects an objective function of stock costs and shortage costs. The static simulation is used to describe deterministic system behavior. In other cases, if the system behavior cannot be predicted deterministically, the behavior results from events over time that influence one another - in such cases, discrete-event simulation is a commonly used [11].

Within optimization algorithms, there are exact procedures that determine an optimum, and optimization heuristics that can determine a good solution for complex practical problems in the practically available computing time [12]. Computation time is especially critical when simulation is used as an evaluation function, as simulation is usually computationally intensive. Metaheuristics are used for practical problems with complex search spaces in which there are many local optima, which are robust to local optima that simple local search procedures, such as hill-climbing procedures, cannot overcome. Rule based heuristics are less universally applicable and require a known optimization strategy, but they are computationally efficient. *Kamhuber et al.* [13] give an example of combining efficient rule-based heuristics, based on human planning expertise, with metaheuristics, in conjunction with discrete-event simulation in production planning. This example demonstrates that a combination or hybridization of the methods can prove to be more useful and advantageous to achieve efficient planning.

In this paper, a heuristic with a static simulation is used as an evaluation function for the selection of the most suitable safety stock calculation. The safety calculation methods are themselves also usually heuristics, that have been established as standard methods in their specific planning domain. Table 1 gives an overview of literature on simulation, optimization and simulation optimization of safety stocks.

Table 1: Literature research on simulation, optimization and simulation-based optimization of safety stocks

Simulation of safety stocks	**Optimization of safety stocks**	**Simulation-based optimization of safety stocks**
Schmidt et al., (2012)	*Gansterer et al.*, (2013)	*Mayer et al.*, (2020)
Gansterer et al., (2013)	*Hernandez-Ruiz*, (2016)	*Claus et al.*, (2018)
Nenni et al., (2013)	*Albrecht*, (2017)	*Bracht et al.*, (2018)
Hernandez-Ruiz, (2016)	*Avci et al.*, (2017)	*Wenzel et al.*, (2017)
Albrecht, (2017)	*Gruler et al.*, (2018)	*Gutenschwager et al.*, (2017)
Avci et al., (2017)	*Ghadimi et al.*, (2020)	*Walmann et al.*, (2016)
Gruler et al., (2018)	*Barrios et al.*, (2020)	*Hanschke*, (2015)
Ghadimi et al., (2020)	*Sourirajan et al.*, (2008)	*Witthaut et al.*, (2015)
Barrios et al., (2020)	*Keskin et al.*, (2015)	
	Schuster-Puga et al., (2016)	
	Park, (2020)	

From the publications in Table 1, the authors highlighted in dark grey were identified as especially relevant publications for this work: *Nenni et al.* [14] evaluate the level of service for safety stocks calculated with different formulas and compare, whether the level of service determined by simulation corresponds to the target level of service of the safety stock level. *Schmidt et al.* [15] deal with the concept of virtual safety stock and evaluate its effectiveness by means of simulation. Freely selected values for lead times, consumption values and their standard deviations function as input data in both works. *Schmidt et al.* use 250 days as the simulation period, which corresponds roughly to the total working days in a year. *Nenni et al.* simulate over 50.000 periods. Due to the uncertainties in demand and replenishment time contained in the models, a single simulation run is not meaningful. For this reason, the simulation results in the papers are arithmetically averaged after 10 or 15 simulation runs. *Nenni et al.* exclusively use the safety stock formula with uncertain lead time (Theory of Constraints), whereas *Schmidt et al.* provide a recommendation matrix for the selection among 9 safety stock algorithms. No real company data is used in each case.

The Paper at hand utilizes 11 safety stock algorithms and provides an evaluation based on a company use-case in the capital goods industry and a final *total landed cost* evaluation is carried out.

4. Development of the safety stock optimization method

4.1 Characterization of the case study

The case study was carried out with the disposition-relevant data of a company from the capital goods industry (production of fittings and valves). The company in the case study is embedded in a corporate group and has about 115 employees, an annual turnover of 22.3 million euro, 346 customers from 51 countries, an annual purchasing volume of 11.5 million Euros, 1.780 active suppliers from 61 countries, and uses an ERP as its central IT system. For the case study, the input files are available in a standardized form from the IT systems and are read in via an interface. The optimization method was implemented in a VBA-based MS Excel tool. The objective was to enable users (materials requirements planners, operational purchasers) to plan optimal safety stock levels independently, without requiring expert knowledge in the areas of optimization and simulation.

4.2 Preliminary ranking of algorithms

As the first step, a preliminary priority ranking of the 11 selected calculation methods was determined, independent of the concrete use case and data set. For this purpose, the capabilities of the safety stock algorithms are compared with the requirements from the uncertainty factors of *Wiendahl* [9] in the calculation of stock of material requirements planning. The result of the prioritization is shown in Table 2 (the algorithms are listed with descending priority). At this stage, this prioritization can be used by application companies – depending on the data availability, the best-ranked method can be chosen by the planner. In this paper, this ranking is only an additional orientation, with the final ranking determined via a simulation evaluation presented in section 4.5.

Table 2: Safety stock procedures and operating principle

	Consideration of relative **relationships**	Consideration of uncertainties of the **delivery quantity**	Consideration of uncertainties in **delivery time**	Consideration of uncertainties in **consumption**	Consideration of uncertainties in the **forecast**	Consideration of **probabilities**
Uncertain lead time (Theory of Constraints)	-	-	+	+	-	+
Safety stock dynamic with target service-level	-	-	+	+	-	+
Safety stock with target service-level	-	+	~	+	-	+
Safety stock with dynamic service-levels	-	-	+	+	-	+
Calculation according to service-level	-	-	~	-	~	+
Calculation by using the A-B-C/X-Y-Z method	+	-	-	+	-	-
Dynamic safety stock method	-	-	+	~	+	-
Calculation with a service-level of 100%	-	+	~	+	-	-
Calculation by means of the rough estimate method	-	-	-	-	-	-
Calculation by *Lagrange* method	+	-	-	-	~	~
Calculation by *Percent-Fill* Method	-	-	-	-	~	~

4.3 Data characterization and data preparation for the use case

In the following section, the procedure of data collection as well as the data structure and results are described. As the first step, all data relevant for the application of the algorithms (stock levels, material, disposition master data, consumption data, etc.) are identified based on the 11 selected safety stock methods (see Figure 3) and obtained from the IT systems of the research partner from the capital goods industry. In the process, a total of seven different files in three different file formats (.xlsx, .csv and .pdf) are combined in the VBA-based MS Excel tool by means of an import logic, sorted by article number, and prepared for further use.

The following Figure 4 shows the data required for the heuristic optimization as well as for the subsequent simulation. In addition to the listed information, the article number is imported for each file for accurate sorting.

Data input for safety stock calculation

Data input 1	**Data input 2**	**Data input 3**	**Data input 4**
– Material number – Base unit – Safety stock – Planned delivery time – Cross-plant material status	– Material number – Net price – Currency – Price unit – Tolerance for underdelivery	– Material number – Order quantity – Delivery date – Quantity delivered – Base unit – Order date	– Material number – Goods receipt date – Material short text – Quantity
Data input 5	**Data input 6**	**Data input 7**	
– Material number – Forecast for the next 12 months – Forecasting method – Relation – Mean forecast error – Standard deviation of the forecast error	– Material number – Chosen method – Order lot sizes	– Units of measurement	

Data input for simulation

Data input 1	**Data input 2**	**Data input 3**
– Safety stocks for calculation methods	– Average lead time – Standard deviation of the lead time	– medium consumption – Standard deviation of consumption

Figure 4: Input data for safety stock calculation and simulation

4.4 Calculation of logistical parameters

After data preparation, relevant logistical parameters for the optimization are calculated from the existing historical and forecast-based input data for each article, as a calculation basis for the safety stock algorithms.

Table 3: Logistical parameters for optimization

Average daily consumption (historical)	Standard deviation lead time (historical)
Average monthly consumption (historical)	Average lot size (historical)
Standard deviation daily consumption (historical)	Maximum monthly consumption (historical)
Standard deviation monthly consumption (historical)	Averaged forecast value (future)
Average lead time (historical)	Standard deviation of the forecast values (future)

4.5 Ranking and selection of safety stock algorithms

Using simulation, the safety stock algorithms are ranked: In the process, a subset of the data set is selected for which all 11 safety stock procedures can be applied for each article (if none or not a substantial share of the dataset is fit for all 11 algorithms, only the supported algorithms are selected and ranked). All algorithms are applied to all articles of the subset (products) and the resulting stock levels are simulated over time (material deliveries and material consumption calculated through a time series analysis) and the article-specific service level is determined for each safety stock procedure. For the simulation, optimal purchase order lot sizes have to be defined – this is achieved via a purchase order lot size optimization method, developed by the authors, which uses a *total landed cost* approach as the objective function of the optimization (all relevant costs are considered) and a deterministic simulation as the evaluation function [16].

Thus, mean lead time, mean consumption, and associated standard deviations of items are imported into the simulation and the warehouse inventory level is simulated for 300 days. The different safety stock levels of the individual algorithms are considered. The algorithms with a combination of the highest resulting service level and a low safety stock level are prioritized. As an example, Figure 5 shows a simulation over 300 days for an item with a mean lead time of 5 days and an associated standard deviation of 3 days. The mean consumption of the example item is 1.000 units per day and the associated standard deviation is 750 units. Figure 5 shows that on days 47, 224 and 242, for example, the stock level would fall to 0 units, meaning that the item would not be available for delivery. The simulation calculates a service level of 96% for this article for a safety stock of 4.829 pieces. The optimal safety stock at a given service level of 98% would be approx. 7.000 pieces, which would be delivered by the procedure with uncertain lead time (Theory of Constraints). Repeating these simulation runs with the subset of the data set plus variation of the items thus led to the decision to prioritize this algorithm.

Figure 5: Simulation of the item-specific stock development

The simulation results are translated into a case and data-set specific ranking of the algorithms. The 11 safety stock algorithms have different data requirements, therefore if the necessary data for certain algorithms is unavailable for an article, the next algorithm is tested. The algorithm with the highest available priority is then chosen for each article.

4.6 Application of safety stock calculation

For the sake of economic viability, safety stocks are not to be created for all articles, but only for those with a significant risk of a shortage, the simulation used in section 4.5 is used here again to assess the item-specific risk. All articles are simulated without setting safety stocks first. The simulation results are then evaluated and articles with risk of stock shortages are identified. Next, the optimal safety stock levels for these articles are calculated with the optimal algorithm chosen before for each article (→ section 4.5). For these articles with safety stocks, the simulation is run anew, this time with the optimal safety stock values, to evaluate the successful avoidance of stock-out.

4.7 Optimizing the heuristics based on the simulation results with optimized purchase order lot sizes and safety stocks

In the evaluation, the items with stock shortages are identified and the optimal safety stock levels calculated in the previous safety stock optimization phase are applied to them - no safety stock is defined for items without shortages. It turned out that, despite optimized safety stocks, a safety stock level that was too low was set for some articles, because in some cases there were still understocking costs in the purchase order lot size simulation. Therefore, the heuristics were revised, and, as an example, the "Theory of Constraints" method was given priority over the "Dynamic service-levels" as the more optimal method. As a result, based on the defined heuristics, only items that are at risk of shortage are suggested or issued an optimal safety stock level.

Lastly, another simulation is carried out for determining the *total landed costs*. After this second fine-tuning, optimized purchase order lot sizes and order times are defined for all articles, as well as optimized safety stocks according to demand. The entire procedure is illustrated in Figure 6.

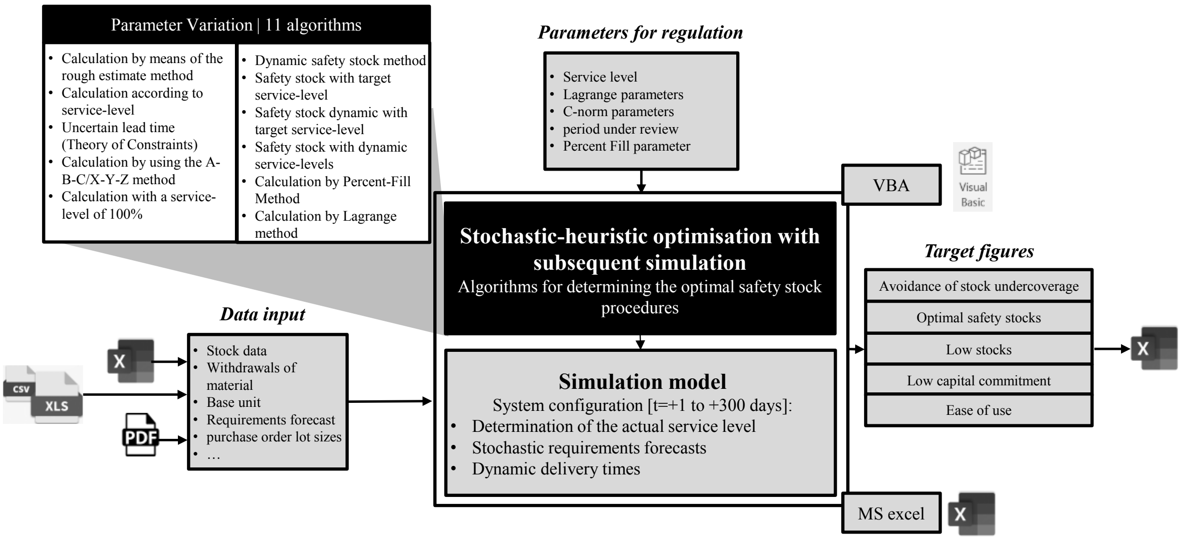


Figure 6: Stochastic-heuristic optimization with subsequent simulation

5. Results and discussion

The results show that using the developed optimization method, different safety stock procedures are identified as optimal for the articles. In this case study, optimal safety stock procedures and safety stocks were determined for 595 articles for which there was a risk of stock shortage, according to the simulation. Of the 11 algorithms considered, 2 were selected by the optimization method. The "Uncertain lead time (Theory of Constraints)" method is used for 76% (452 articles) and the "Safety stock with target service-level" method for 24% (143 articles). The other 9 algorithms have not been selected for this dataset. If the database were expanded from the 595 articles in the application example, other procedures would also be selected, based on the data availability of each article. However, since there is no danger of a stock shortage for these articles, no safety stock was suggested or determined according to the developed optimization method.

The combined consideration of calculation of stock (safety stock method) and purchase order lot size calculation (purchase order calculation method) increases the service level (availability of goods). For all simulated 595 articles with understocking costs (out of 4.066 articles in total), a stock-out could be avoided for almost all articles (476 articles). For 119 articles, a stock-out could not be avoided due to the nature of the initial state at the beginning of the planning period: For those articles, errors in the material requirements planning in the period before the considered planning period have led to a stock level of zero at the start of the planning period, which led to unfulfilled demand right after the start. In principle, this phenomena cannot be avoided.

It must be noted that the increase in the availability of goods through safety stocks is at the expense of warehouse and capital commitment costs, as shown in Table 4. Only costs represented through figures, data and facts in financial accounting records were able to be evaluated. For example, a possible customer fluctuation due to insufficient delivery capability could not be evaluated financially. Through this targeted and optimal application of safety stocks, it was possible to guarantee the service level and the associated availability of goods for those articles with understocking costs (apart from those 119 articles) with a minimal increase in costs of ~0.6 percentage points.

After applying the last (2nd) phase of fine-tuning, the combined optimization of purchase order lot sizes and safety stocks, the *total landed cost* shows higher total costs for this combination compared to a procurement optimized only for purchase order lot sizes, without optimized safety stocks (see Table 4). From the point of view of the *total landed cost* objective function, the optimization result has thus even slightly deteriorated due to the combined optimization, while it has improved from the point of view of the safety stock optimization objective function (service level 99,9% with the lowest possible safety stocks). Since the *total landed cost* approach only considers actual costs incurred and does not consider, for example, the negative effects of a stock-out on customers who could reorient themselves to other suppliers, it is reasonable to suggest refining the objective function from an overall optimization point of view - i.e., the entire material requirements planning.

Table 4: Total landed cost (TLC) consideration before and after optimization of safety stocks

	TLC consideration **before** optimization of the safety stock level	TLC consideration **after** optimization of the safety stock level
Stock shortage costs	15.691 EUR	6.811 EUR
Storage costs	60.855 EUR	62.543 EUR
Capital commitment costs	329.794 EUR	339.317 EUR
Total landed cost (TLC)	406.340 EUR	408.672 EUR

6. Conclusion and outlook

The method was developed using an extensive case study and data set from the capital goods industry. It is based on established calculation methods for purchase orders and safety stocks. In principle, it is therefore suitable for most companies that operate complex material requirements planning. The benefits increase with an increasing number of articles as well as risk factors and other complexity drivers, all of which can be found in the capital goods industry – this is where the digital (partial) automation of planning can prove to be most valuable.

The results show that with the targeted use of digitalization in the calculation of stock of consumption-controlled material requirements planning, the service levels can be significantly improved. In addition, the interaction and interdependence of the main disciplines of calculation of stock and purchase order lot size calculation in materials disposition is also presented. The relevance of data quality and structure in companies was also demonstrated while the study was underway.

Further research work will be aimed at the development of an integrated material requirements planning method, comprising requirements planning, calculation of stock and purchase order lot size calculation. This will include investigating the hierarchy between planning goals and working towards a less sequential planning process in the interest of pursuing an aligned material requirements planning optimization.

Acknowledgements

This research was funded by the *Austrian Research Promotion Agency* (FFG), research promotion 872787. The authors would like to thank all case study partners for their contribution.

References

[1] Paulus-Rohmer, D., Schatton, H., Bauernhansl, T., 2016. Ecosystems, Strategy and Business Models in the age of Digitization - How the Manufacturing Industry is Going to Change its Logic. Procedia CIRP 57, 8–13.

[2] Arnolds, H., Heege, F., Röh, C., Tussing, W., 2016. Materialwirtschaft und Einkauf. Springer Fachmedien Wiesbaden, Wiesbaden, 474 pp.

[3] Jacob, M., 2013. Management und Informationstechnik: Eine kompakte Darstellung. Springer Fachmedien Wiesbaden; Imprint: Springer Vieweg, Wiesbaden, 187 pp.

[4] Wannenwetsch, H., 2014. Integrierte Materialwirtschaft, Logistik und Beschaffung. Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Sihn, W., 2016. Produktion und Qualität; Organisation, Management, Prozesse.

[6] Minner, S., Stößlein, M., 2015. Neuere Ansätze und Methoden zur Festlegung von Sicherheitsbeständen, in: Claus, T., Herrmann, F., Manitz, M. (Eds.), Produktionsplanung und –steuerung. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 209–224.

[7] Wischmann, S., Hartmann, E.A. (Eds.), 2018. Zukunft der Arbeit – Eine praxisnahe Betrachtung. Springer Berlin Heidelberg, Berlin, Heidelberg.

[8] Peffers, K., Tuunanen, T., Rothenberger, M.A., Chatterjee, S., 2007. A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24 (3), 45–77.

[9] Wiendahl, H.-P., Wiendahl, H.-H., 2020. Betriebsorganisation für Ingenieure.

[10]Pfohl, H.-C., 2018. Logistiksysteme. Springer Berlin Heidelberg, Berlin, Heidelberg.

[11]März, L., Krug, W., Rose, O., Weigert, G. (Eds.), 2011. Simulation und Optimierung in Produktion und Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[12]Jourdan, L., Basseur, M., Talbi, E.-G., 2009. Hybridizing exact methods and metaheuristics: A taxonomy. European Journal of Operational Research 199 (3), 620–629.

[13]Kamhuber, F., Sobottka, T., Heinzl, B., Sihn, W., 2019. An Efficient Multi-Objective Hybrid Simheuristic Approach for Advanced Rolling Horizon Production Planning, 2108–2118.

[14]Nenni, M.E., Schiraldi, M.M., 2013. Validating Virtual Safety Stock Effectiveness through Simulation. International Journal of Engineering Business Management 5, 41.

[15]Schmidt, M., Hartmann, W., Nyhuis, P., 2012. Simulation based comparison of safety-stock calculation methods. CIRP Annals 61 (1), 403–406.

[16]Schmid, A., Sobottka, T., Lielacher, M., Sihn, W., 2021. Simulationsbasierte Optimierung von Bestelllosgrößen in der verbrauchsgesteuerten Materialdisposition der Investitionsgüterindustrie. 19. ASIM Fachtagung - Simulation in Produktion und Logistik.

Biography

After graduating in 2007, **Alexander Schmid** (*1977) developed and was responsible for the department of supply chain management at *Knorr-Bremse GmbH*. Since Oct. 2013, Dipl.-Ing. Alexander Schmid has overseen various teaching activities in the areas of supply chain management and logistics at the *Vienna University of Technology* and is employed as a research assistant at *Fraunhofer Austria.*

Thomas Sobottka (*1985) finished his Phd in 2017 at the Vienna University of Technology. In his current research, he supervises projects for the simulation-based energy-aware optimization of production processes, production planning and control in short to long-term planning horizons and digital-twin applications in production planning and control.

Wilfried Sihn (*1955) Prof. Dr. Wilfried Sihn has been active in the field of applied Research for more than 30 years, taking part in more than 300 industrial projects. His areas of expertise include production management, corporate organization, enterprise logistics, factory planning, order management, life-cycle management, maintenance, modelling and simulation, and business process reengineering.

3rd Conference on Production Systems and Logistics

Modular, Digital Shopfloor Management Model – A Maturity Assessment For A Human-Oriented Transformation Process

Magnus Kandler[1], Philip Gabriel[1], Vincent Schröttle[1], Marvin Carl May[1], Gisela Lanza[1]

[1]Karlsruhe Institute of Technology, wbk – Institute of Production Science, Kaiserstr. 12, Karlsruhe, Germany

Abstract

Currently digitization and Industry 4.0 are some of the most important trends influencing production processes and companies. In a lot of companies Shopfloor Management (SFM) is used to manage production and sustain as well as grow a strategic leadership through lean and efficient processes. Its main goal is to bring focus back to the shopfloor, where value-adding processes are performed, and to increase the connection between managers and shopfloor workers. In light of the aforementioned trend towards continued digitization, this produces a field of tension in which a trend enabling increasingly remote control of production through new digital solutions meets a principle focusing on being on site as well as intensive communication. Combing both aspects can prove to be complex as can be seen by studies showing a wide distribution of SFM in a selection of German production companies, but with exceptionally low usage of digital technology.

A significant cause of this is the overwhelming number of possible aspects of Industry 4.0 increasing the difficulty of selecting manageable ideas, aggravated by a lack of integration of necessary employee and organizational change. Purpose of this paper is, therefore, to propose a theoretical basis for a digitization of SFM through the development of a model including disjunct aspects of SFM as well as a maturity index providing a means to discern different levels of digital solutions. To provide a practical viewpoint in addition to the theoretical basis, concrete SFM methods are collected and mapped to fitting areas as well as maturity levels of the underlying model. Additionally, a two-step questionnaire is conceived to allow for a selective proposition of said methods to derive an implementation sequence. All aspects of the approach are finally validated with an example company.

Keywords

Digital Shopfloor Management; Lean Leadership; Human-oriented Shopfloor Management

DOI: https://doi.org/10.15488/12153

1. Introduction and Motivation

Current trends like globalization, individualization and shorter product life-cycle times are causing qualities like adaptability and flexibility to become more and more important [1]. Highly efficient and flexible production systems with a flexible organisation as well as satisfied and motivated employees are required. However, in a survey conducted by Staufen AG in 2019, 70% of companies state that changes are predominantly imposed from above and only 43% of respondents state that their organizational structures are designed to be flexible and changeable [2]. This contradicts the vision of decentralized design and optimization as well as changeable companies and at the same time leads to a low acceptance of the methods used in the production system [3]. Technology is thus no longer the limiting factor, but human mutability [1]. Besides technological changes such as plug-and-produce or self-organized production, Shopfloor Management (SFM), hence, is a key enabler of a flexible and adaptable company, for example through allowing increased transparency on the shopfloor [4]. In particular, it supports short-cycle decisions that are associated with organizational and human adaptability.

Because of the progressive digitization and increasing implementation of Industry 4.0 the concept of SFM faces a large number of changes, creating new potentials. These are, however, often accompanied by concerns of employees [5]. At the same time, digital networks also possess inherent risks for example in the form of data security, long latency periods and information overload [6]. Additionally, due to the many ways in which Industry 4.0 is currently affecting companies, it must be analysed how SFM can be meaningfully expanded through digitization in order to promote decentralized decisions in production [7,8].

The goal of this paper is to provide an approach to combine Industry 4.0 and SFM including all affected dimensions (technological, human and organizational) in change. Therefore, in section 2 available literature concerning a theoretical basis for the digitization of SFM is analysed. Based on the existing models the modular digital SFM model including a maturity index is presented in sections 3.1 and 3.2. To expand the theoretical view towards a real-world application, concrete digital methods are then added in section 3.3. To allow for a further customization of the implementation process a two-step questionnaire is presented in section 3.4 enabling the allocation of a company within the theoretical model and deriving suitable steps towards increasing the digital maturity level. In section 4 the approach is exemplary applied to a production company and given recommendations for a digital SFM transformation are shown. Finally, the approach is discussed and potential for future research is shown.

2. State of the art

In order to allow for a structured evaluation of available research towards (digital) SFM as well as implementation approaches, a set of criteria is defined as a first step. The first one is the availability of a method tool box (C1). This is necessary to narrow down the immense amount of possible solutions within Industry 4.0 and build a basis for a concrete plan of steps to integrate digital solutions into SFM [9]. This also allows to fully integrate the peculiarity of digital change processes. The second criterion is the integration of concrete aid towards implementation of (digital) methods (C2). As mentioned before, digital solutions not only come with technical challenges, but implementation also has to consider organizational as well as human challenges, which ought to be considered to fulfil C2 [10]. Criterion C3, the availability of an assessment (tool), aims a supplying a means to generate a, as much as possible, reproduceable strategy towards implementation for companies instead of being limited to a theoretical view. Approaches should also incorporate a multi-level maturity index (C4). This allows to separate different levels of digital solutions whilst still placing them along a linear guideline steering the digitization [11].

The first type of identified literature are analogue SFM models. An exemplary approach is provided by Kiyoshi Suzaki, which describes an extensive basis for the transformation of a company from a “traditional” way of management to a “modern” SFM [12]. By using examples and experiences a simple implementation approach is provided. He, however, only describes the change necessary for each single aspect of production as well as the ideal result, whereas a concrete implementation strategy is only slightly addressed. Another comparable approach is chosen by [13]. In addition to both an overview over methods of SFM as well as detailed explanation for each aspect, Conrad et al. also facilitate a concrete SFM implementation by embedding the methods in a step-based implementation strategy with clear instructions. As a further aid multiple concrete design examples of different methods are given to be used as inspiration for custom implementations. All approaches, however, lack an integration of digital change in SFM. All of the identified approaches for analogue SFM mainly address C1 and C2 by providing different (analogue) methods to implement SFM as well as addressing necessary changes in the human and organizational perspective. None, however, feature an assessment tool or, due to their analogue focus, a maturity index.

The second type of identified literature addresses digital SFM. These can be divided into approaches being based on analogue models as well as newly developed models. Brenner, for example, starts off with a description of analogue SFM [14]. He then uses these parts as an agenda to demonstrate different exemplary ways to digitize processes. A similar approach is chosen by Meißner et al. [15,16]. Building upon their existing SFM model, they first define a target state and thereby enhance the analogue model with digital aspects that support and ease the use of it. Additionally, they propose traditional change management models for strategic change combined with tailored recommendations as a way to introduce digitalization to the shopfloor. A different approach is chosen by [17] as well as [18]. Instead of starting with an existing SFM model, Rauch et al. develop a description of a final goal stage concerning the digital abilities for SFM [17]. They then describe the different methods having to be implemented to achieve the final state. Bock et al. [18] also design their approach without an analogue model, but instead define four different maturity levels for digital SFM ranging from analogue to autonomous. Instead of proposing defined steps for achieving the target state, they offer an extensive description for them. This in turn allows them to be used as a form of vision, for which to achieve individual steps need to be taken (like different software modules or functions). Most of the identified approaches not only address technical challenges, but integrate all necessary aspects with few authors also defining (basic) maturity indexes (e.g. [4]). Similarly, only a minority of the identified literature offers concrete methods for achieving the proposed visions (C1) and in turn also assessment tools to derive methods to be implemented (C3).

A third type of relevant literature like the works of Liebrecht et al. [19] and the practical framework for SFM implementation by Hartner et al. [6] are about implementation procedures for digital solutions. As can be seen by the recommendation of [16] to implement digital SFM with strategic change management methods, digitization often means fundamental change for employees and companies in general. As [8] mention, technical changes are often more advanced compared to social and organisational change processes. To overcome this inequality, they propose an iterative framework for implementing dSFM integrating the technical, organization and human dimension. They thereby add a human-centered perspective to the predominating technical views of digital SFM. All approaches in this area, however, lack concrete methods for SFM and only propose general digital methods, if at all (C1). Yet, they sometimes offer maturity indexes (like [11]) and mostly address the necessary change in the organizational as well as human dimension (C4 and C2)

As can be seen in Figure 1, there is no current research activity towards a complete approach fulfilling all defined criteria and enabling a step-by-step processes to digitize SFM based on giving concrete method recommendations. Therefore, we present a modular SFM model, which supports decentralized decisions like process optimization or short-time production steering. This model includes maturity index as well as a toolbox that presents dSFM methods for each level of a digital SFM, combining the theoretical viewpoint given with the underlying model with real-world application through concrete methods, that can be implemented in companies. The approach is further enhanced in regards to real-world applicability by a human-oriented dSFM maturity assessment, which calculates the actual stage of dSFM in a company and therefore determines an individual development process for the company towards a target dSFM-stage based on the proposed methods located in the theoretical model.

● Fully fulfilled
◑ Partially
○ Not fulfilled

Research Approaches Criteria	[12]: Suzaki	[13]: Conrad et al.	[14]: Brenner	[15], [16]: Meißner et al.	[17]: Rauch et al.	[18]: Bock et al.	[4]: Lanza et al.	[19]: Liebrecht et al.	[6]: Hartner et al.	[8]: Kandler et al.	[11]: Schuh et al.
Method tool box (C1)	◑	◑	●	◑	●	◑	◑	◑	◑	◑	◑
Concrete aid towards implementation (C2)	●	◑	◑	◑	○	○	◑	●	○	●	◑
Assessment tool (C3)	○	○	○	●	○	○	○	○	○	◑	○
Multi-level maturity index (C4)	○	○	○	○	○	●	●	○	○	○	●

Figure 1: Overview of identified literature (Picture by authors)

3. Methodology

3.1 Development of the underlying model

As suitable model is the necessary basis for the assessment of the current state as well as the vision and, thereby, allows to derive the next steps towards improving maturity [20]. However, as can be seen in the previous chapter, many theoretical models only feature a limited scope, either missing human and /or organizational perspective or only partially addressing the challenges stemming from digitization of SFM.

Therefore, the initial step towards conceiving an approach for achieving digital SFM is the development of an adequate model. For this, the concept of SFM by [4] is used as a basis to be evolved. Analogue SFM already is not only a collection of methods to be implemented, but also includes a wider view with fundamental change to organizational structure enabling necessary change for successful SFM processes [12]. Additionally, the human perspective also plays a big part for achieving a functioning SFM as well as digitization in general [21,22]. Both perspectives are already rudimentarily featured in the underlying model by [4]. It is, however, lacking the integration of digital aspects at its core. To address this and following the approach by [19], all six areas are separated into "enabler"- as well as "potential"-categories. The enablers include areas allowing new potentials to be achieved, whereas the "potentials" offer a competitive advantage when implemented. Accurate data is the basis not only of digitization, but also of analogue SFM [23,12]. When digitizing data, however, many aspects have to be considered like data security, the distribution of automatic vs. manual data collection as well as efficient storage [24–26]. These aspects are integrated into the model by adding the separate **data** area. Another important enabler-area is **IT-enablers**, which contains the technical basis of digitization. This also has to be integrated separately as there are many different possibilities of achieving for example widespread networking of digital systems. Yet, old infrastructure, varying standards etc. can pose big risks for safety, scalability or other important future success factors [9,15]. The last enabling area is **Key Performance Indicators** or **KPI** which is already used by [4]. These provide the main instrument to achieve transparency on the shopfloor, an important aspect allowing all employees to participate and ease the identification of malfunctions [27,12]. In the developed model, the define the gateway between data and usable information for the potential areas.

To keep the model compact, for the "potential"-categories, some of the dimensions of the model of [4] are grouped together. One area is thus defined by combining **Meetings & Knowledge Exchange**. Meetings play an important role in structuring communication and as a basis for other processes in SFM [4,12]. The digitization of this area has to be carefully balanced to not lose focus on the shopfloor for example by employing digital meetings to replace regular presence [6]. To allow for the integration of all employees in SFM, a continued qualification and exchange of knowledge is necessary [13,12]. This will also play a big role in successful digitization and offers vast possibilities through digital evolvement [28,29]. The area of **Measures & Problem-Solving** integrates another important part of SFM. Through continuous identification and solving of problems and realizing improvement potentials in production processes, its efficiency can be increased. This plays an important role in the ability of SFM to secure and increase competitive advantage. Through new possibilities of data processing etc. in this area, these processes can be significantly enhanced, but also face problems through a risk of lacking integration of employees in automatic processes and therefore declining readiness in participating in continuous production improvement [14]. Participation of employees however will also play an important role in future problem solving [30]. The last area of the model is defined as **Resource Control**.

As described by Ganschar et al., the human will still be the centre of the modern fabric [31]. Thus, employee perspective acts as a base for the model, being represented by necessary qualifications as well as acceptance. To fully reflect necessary changes in the organizational dimension, the model is supplemented by the areas of organizational guidelines as well as lean leadership. The complete model with all areas and dimensions (see 3.3) can be seen in Figure 2.

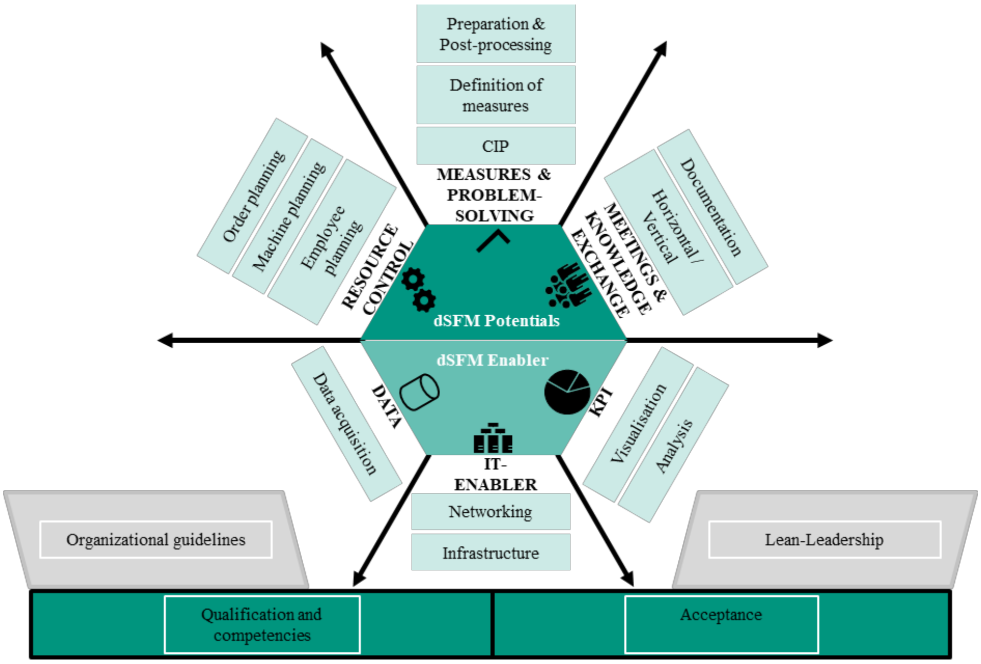


Figure 2: Developed SFM model already with dimensions described in 3.3 (Picture by authors)

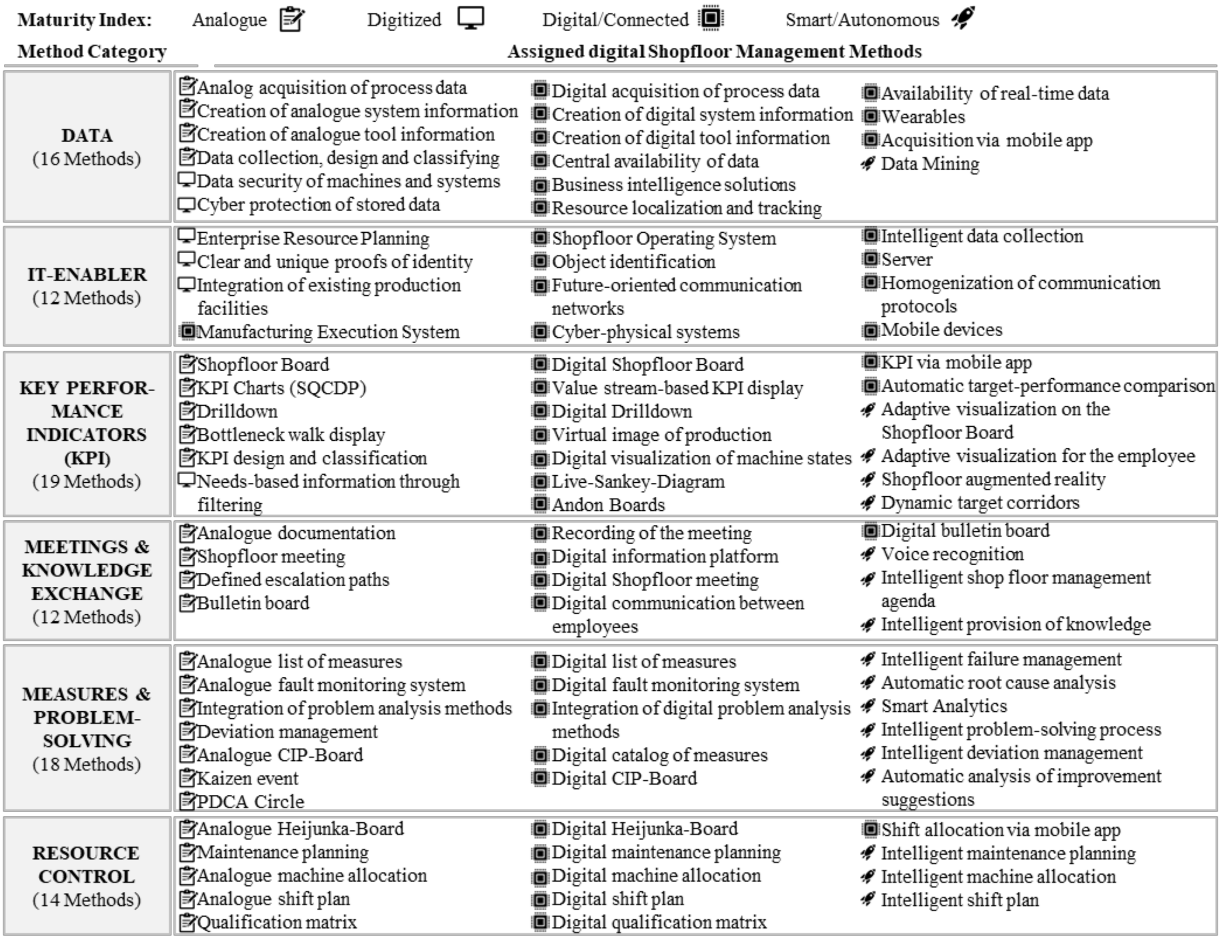
3.2 Development of a digital SFM Maturity Model

As an addition to the developed model, a suitable maturity index is conceived, which is based on the study of [4]. The matrix by [4] containing the three criteria (I4.0 levels as defined by [11], real-time capabilities and level from analogue to automated) is thus transformed into a linear model. Therefore, the basic distinction between analogue, digital and automated is used and slightly renamed. Since certain methods warranted a closer distinction between digital and analogue, a fourth level called “digitized” is introduced to form the final maturity index [15]. “Analogue” includes methods without any kind of digital support, whereas “digitized” is mostly used for methods with basic digital support (like digital lists). “Digital /

connected” includes methods with complex digital support and connective features, but only “smart / autonomous” contains methods with extensive digital support, where algorithms etc. gain partial decision competence and / or forms of intellect.

3.3 Collection and allocation of concrete (digital) SFM methods

SFM according to the definition by Liebrecht et al. is an independent method or management tool [32], whereas in the context of the Lean-Philosophy SFM is described as an understanding of leadership [23]. Addressing the inhomogeneous definition of SFM and to address criterion C1 a major structured literature search is performed to identify existing methods of (digital) SFM. The research is focusing on already existing methods for the SFM but includes ideas and visions as well. Therefore, the identified methods range from being widely implemented already in lean companies (like Shopfloor Boards used to visualize KPI’s [23]) to being highly innovative and thus not having been implemented in most companies yet, like trend analytics to predict deviations of KPI’s [9]. A complete overview of all identified methods can be seen in Figure 3.



Maturity Index: Analogue 📝 Digitized 🖵 Digital/Connected ▣ Smart/Autonomous 🚀

Method Category	**Assigned digital Shopfloor Management Methods**		
DATA (16 Methods)	📝Analog acquisition of process data 📝Creation of analogue system information 📝Creation of analogue tool information 📝Data collection, design and classifying 🖵Data security of machines and systems 🖵Cyber protection of stored data	▣Digital acquisition of process data ▣Creation of digital system information ▣Creation of digital tool information ▣Central availability of data ▣Business intelligence solutions ▣Resource localization and tracking	▣Availability of real-time data ▣Wearables ▣Acquisition via mobile app 🚀 Data Mining
IT-ENABLER (12 Methods)	🖵Enterprise Resource Planning 🖵Clear and unique proofs of identity 🖵Integration of existing production facilities ▣Manufacturing Execution System	▣Shopfloor Operating System ▣Object identification ▣Future-oriented communication networks ▣Cyber-physical systems	▣Intelligent data collection ▣Server ▣Homogenization of communication protocols ▣Mobile devices
KEY PERFOR-MANCE INDICATORS (KPI) (19 Methods)	📝Shopfloor Board 📝KPI Charts (SQCDP) 📝Drilldown 📝Bottleneck walk display 📝KPI design and classification 🖵Needs-based information through filtering	▣Digital Shopfloor Board ▣Value stream-based KPI display ▣Digital Drilldown ▣Virtual image of production ▣Digital visualization of machine states ▣Live-Sankey-Diagram ▣Andon Boards	▣KPI via mobile app ▣Automatic target-performance comparison 🚀 Adaptive visualization on the Shopfloor Board 🚀 Adaptive visualization for the employee 🚀 Shopfloor augmented reality 🚀 Dynamic target corridors
MEETINGS & KNOWLEDGE EXCHANGE (12 Methods)	📝Analogue documentation 📝Shopfloor meeting 📝Defined escalation paths 📝Bulletin board	▣Recording of the meeting ▣Digital information platform ▣Digital Shopfloor meeting ▣Digital communication between employees	▣Digital bulletin board 🚀 Voice recognition 🚀 Intelligent shop floor management agenda 🚀 Intelligent provision of knowledge
MEASURES & PROBLEM-SOLVING (18 Methods)	📝Analogue list of measures 📝Analogue fault monitoring system 📝Integration of problem analysis methods 📝Deviation management 📝Analogue CIP-Board 📝Kaizen event 📝PDCA Circle	▣Digital list of measures ▣Digital fault monitoring system ▣Integration of digital problem analysis methods ▣Digital catalog of measures ▣Digital CIP-Board	🚀 Intelligent failure management 🚀 Automatic root cause analysis 🚀 Smart Analytics 🚀 Intelligent problem-solving process 🚀 Intelligent deviation management 🚀 Automatic analysis of improvement suggestions
RESOURCE CONTROL (14 Methods)	📝Analogue Heijunka-Board 📝Maintenance planning 📝Analogue machine allocation 📝Analogue shift plan 📝Qualification matrix	▣Digital Heijunka-Board ▣Digital maintenance planning ▣Digital machine allocation ▣Digital shift plan ▣Digital qualification matrix	▣Shift allocation via mobile app 🚀 Intelligent maintenance planning 🚀 Intelligent machine allocation 🚀 Intelligent shift plan

Figure 3: Overview of identified methods and corresponding area and maturity level (Picture by authors)

Methods from the research period having been classified relevant are then transformed into standardized profiles, which enable clear and fast summary of the elements necessary for a successful implementation and application of the shopfloor method. These profiles are summarized in a modular SFM toolbox similar to Liebrechts Industry 4.0 toolbox [32]. The profiles contain various information in regards to expected effects and benefits for SFM and production in general, but also surrounding effects like employee reaction / necessary qualification as well as needed organizational adaption. This addresses the criterion C2 by enhancing the method description with organizational and human perspective. The methods are then initially mapped to the most appropriate model area and maturity level. To allow for an even more precise allocation, the areas are separated into dimensions like "Documentation" and "Horizontal / Vertical" for the category "Meetings & Knowledge Exchange" based on a clustering of contained methods.

3.4 Development of a digital SFM Maturity Assessment

In order to allow the assessment of a company's current level of maturity concerning digital SFM as well as rate their vision, a short as well as detailed questionnaire is developed following the approach of Schumacher et al [33]. Their work is an ideal basis as it is a combination of numerous maturity assessment concepts and is easily transferred to SFM and its digitization. Instead of a single survey, however, in our approach two different questionnaires are used. This two-level approach allows interested companies to answer a short Quick Check and receive more general recommendations including suitable methods to be implemented as a rapid way to develop towards digital SFM. Users needing an in-depth guidance can then answer detailed questions regarding the previously identified lacking area(s) of SFM.

As a first step, a description for the ideal final state of each tuple of maturity level and field of application is created based on the profiles as well as additional literature research. The final states consist of a combination of possibilities offered by the methods in the respective tuple as well as their main advantages when reached. The conceived final stages are then in turn analysed for their most important contents and afterwards transformed into items for the Quick Check based on available literature concerning questionnaire best practices like [34]. For the development of the more detailed questionnaire allowing a more exact rating of the current maturity level, the method profile is used as a source of information and in turn transformed into one or more items describing each aspect of a given method. Closely related methods are as much as possible combined into single items to limit the size and in turn answering time of the questionnaire. The complete process is visualized in Table 1.

Table 1: Example for process from methods to item used to derive items of maturity assessment for the Category of Data Acquisition and Maturity Level Analogue

Step	Content
Input: Methods	Analogue Acquisition of Process Data, Creation of analogue machine datasheets, Creation of analogue tool datasheets, Data Acquisition Design and Classification
Throughput: Ideal final state	Up-to-date information through short cyclical data acquisition, Identification of relevant data, Structured data form and acquisition process
Output: Derived item	Do rules for the classification and structuring of data exist on the shopfloor?

To allow for a comparison of results, answering options were mapped to corresponding maturity levels in the Quick Check and in turn recommendations. For the detailed questions, answering options were mapped to distinct methods or method implementation progress. To finalize the surveys, both questionnaires are then pre-tested. To make sure, that each question is understandable, three different personae are conceived representing different types of users to be expected based on surveys for example from [35]. Main points of each persona are their knowledge about digitalization as well as SFM, which are used to make sure each question is understood the correct way and addresses the correct underlying aspect or method.

4. Application and verification

The presented digital SFM quick check has been applied with an international production company producing measuring devices. It is characterized by a decentralized organizational structure in production with autonomous assembly teams and it's at the beginning of implementing Industry 4.0. The application was carried out in three sequential steps: In the first step the partner answered the Quick Check consisting of approx. 46 questions, second the calculated maturity levels were discussed in semi-structured interviews with an expert of the company and third the levels were adapted where necessary. The answers given and the calculated maturity levels (see Figure 4) in the Quick Check then fit the described situation of the company.

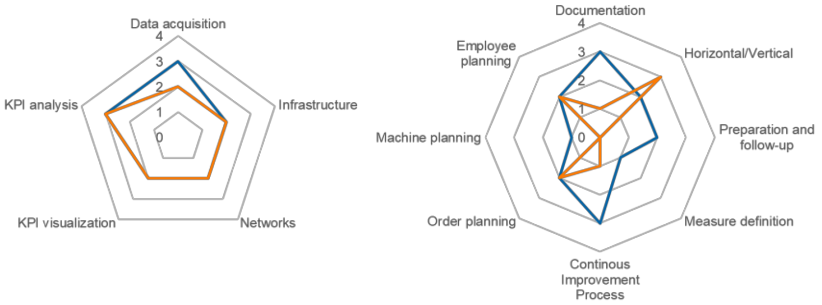


Figure 4: Resulting maturity levels for enabler (left) and potential dimensions (right) from the Quick Check visualized using a Microsoft® Excel graph (blue = target level, orange = current level, Picture by authors)

As being identified in the Quick Check as lacking, the category of Preparation and follow-up was thoroughly examined using the detailed questionnaire. There, the low level was confirmed and it was found, that necessary analogue requirements in the form of structured problem collection and measure documentation was missing. Due to their otherwise comparably high level in the enabler areas, a digital fault monitoring system as well as a digital measure list were recommended to be implemented.

5. Conclusion

The presented modular digital SFM model summarizes single SFM methods. Because of the described maturity levels and the developed assessment for each category the presented approach is a suitable tool for supporting the company-specific (digital) SFM implementation. The exemplified application of the model showed the generally possibility of estimation of the digital maturity. However, some of the single questions need to be reworked, so that a self-assessment will be practicable.

In further research the guidelines of the digital SFM model should be concretised with the help of empirical research methods like item tests and regression analysis. For example, the relevant competences and qualifications must be defined for each employee role in SFM. The corresponding competencies must be available among the employees so that successful SFM processes can be realized. If the maturity assessment is combined with an employee self-evaluation of the human-oriented target system [8] a benchmark study is facilitated, which can be used to determine the success factors of SFM in terms of high employee acceptance and employee productivity.

Acknowledgements

The authors gratefully acknowledge funding from “teamIn”, which is funded by the Federal Ministry of Education and Research and the European Social Fund as part of the “Future of Work” program (funding reference number: 02L18A140) and supervised by the Project Management Agency Karlsruhe (PTKA).

References

[1] Lanza, G., Nyhuis, P., Fisel, J., Jacob, A., Nielsen, L., Schmidt, M., Stricker, N., 2018. Wandlungsfähige, menschzentrierte Strukturen in Fabriken und Netzwerken der Industrie 4.0. Herbert Utz Verlag, München.

[2] Staufen AG, 2019. German Industry 4.0 Index 2019: A study from Staufen AG and Staufen Digital Neonex GmbH.

[3] Abel, J., Hirsch-Kreinsen, H., Wienzek, T., 2019. Acceptance of Industry 4.0: Final Report on an Exploratory Empirical Study of German Industry. Plattform Industrie 4.0.

[4] Lanza, G., Hofmann, C., Stricker, N., Biehl, E., Braun, Y., 2018. Auf dem Weg zum digitalen Shopfloor Management: Eine Studie zum Stand der Echtzeitentscheidungsfähigkeit und des Industrie 4.0-Reifegrads. Studie. Karlsruher Institut für Technologie, Karlsruhe. https://www.wbk.kit.edu/downloads/Auf%20dem%20Weg%20zum%20digitalen%20Shopfloor%20Management.pdf.

[5] Parry, G.C., Turner, C.E., 2006. Application of lean visual process management tools. Production Planning & Control 17 (01), 77–86.

[6] Hartner, R., Mezhuyev, V., Tschandl, M., Bischof, C., 2020. Digital Shop Floor Management, in: Proceedings of the 2020 9th International Conference on Software and Computer Applications. ICSCA 2020: 2020 9th International Conference on Software and Computer Applications, Langkawi Malaysia. Association for Computing Machinery, New York,NY,United States, pp. 41–45.

[7] Hertle, C., Siedelhofer, C., Metternich, J., Abele, E., 2015. The next generation shop floor management – how to continuously develop competencies in manufacturing environments. The 23rd International Conference on Production Research.

[8] Kandler, M., May, M.C., Kurtz, J., Kuhnle, A., Lanza, G., 2022. Development of a Human-Centered Implementation Strategy for Industry 4.0 Exemplified by Digital Shopfloor Management, in: Andersen, A.-L., Andersen, R., Brunoe, T.D., Larsen, M.S.S., Nielsen, K., Napoleone, A., Kjeldgaard, S. (Eds.), Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems. Springer International Publishing, Cham, pp. 738–745.

[9] Kagermann, H., Wahlster, W., Helbig, J., 2013. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0.

[10] Schlicher, K.D., Paruzel, A., Steinmann, B., Maier, G.W., 2020. Change Management für die Einführung digitaler Arbeitswelten, in: Maier, G.W., Engels, G., Steffen, E. (Eds.), Handbuch Gestaltung digitaler und vernetzter Arbeitswelten. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 347–382.

[11] Schuh, G., Anderl, R., Gausemeier, J., Hompel, M. ten, Wahlster, W., 2020. Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten.

[12] Suzaki, K., 1994. Die ungenutzten Potentiale: Neues Management im Produktionsbetrieb. Hanser, 392 pp.

[13] Conrad, R.W., Eisele, O., Lennings, F., 2019. Shopfloor-Management - Potenziale mit einfachen Mitteln erschließen: Erfolgreiche Einführung und Nutzung auch in kleinen und mittelständischen Unternehmen, 1. Aufl. 2019 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 70 pp.

[14] Brenner, J., 2019. Shopfloor Management und seine digitale Transformation: Die besten Werkzeuge in 45 Beispielen : mit 116 Bildern und 2 Tabellen. Hanser, München, 246 pp.

[15] Meißner, A., Grunert, F., Metternich, J., 2020. Digital shop floor management: A target state. Procedia CIRP 93, 311–315.

[16] Meißner, A., Hertle, C., Metternich, J., 2018. Digitales Shopfloor Management – Ihr Weg zur vernetzten Fabrik. Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (5), 281–284.

[17] Rauch, E., Rojas, R., Dallasega, P., Matt, D.T., 2018. Smart Shopfloor Management. Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (1-2), 17–21.

[18] Bock, T., Höfer, S., 2021. Autonomisierung von Shopfloor Management. Zeitschrift für wirtschaftlichen Fabrikbetrieb 116 (3), 139–143.

[19] Liebrecht, C., 2020. Entscheidungsunterstützung für den Industrie 4.0 - Methodeneinsatz. Dissertation.

[20] Becker, J., Knackstedt, R., Pöppelbuß, J., 2009. Developing Maturity Models for IT Management. Bus. Inf. Syst. Eng. 1 (3), 213–222.

[21] Malessa, N., Ast, J., Kandler, M., Ströhlein, K., Nyhuis, P., Lanza, G., Nieken, P., 2020. Digitale Führung und Technologien für die Teaminteraktion von morgen. Zeitschrift für wirtschaftlichen Fabrikbetrieb 115 (7-8), 540–544.

[22] Ullrich, A., Vladova, G., Thim, C., Gronau, N., 2019. Organisationaler Wandel und Mitarbeiterakzeptanz. Vorgehen und Handlungsempfehlungen, in: Obermaier, R. (Ed.), Handbuch Industrie 4.0 und digitale Transformation. Betriebswirtschaftliche, technische und rechtliche Herausforderungen. Springer Gabler, Wiesbaden, Heidelberg, pp. 565–587.

[23] Bertagnolli, F., 2020. Lean Management. Springer Fachmedien Wiesbaden, Wiesbaden.

[24] Dietrich, M., 2021. Digitales Shopfloor Management in SAP-Systemumgebungen. Springer Fachmedien Wiesbaden, Wiesbaden.

[25] Wölfl, S., Leischnig, A., Ivens, B., Hein, D., 2019. From Big Data to Smart Data – Problemfelder der systematischen Nutzung von Daten in Unternehmen, in: Becker, W., Eierle, B., Fliaster, A., Ivens, B., Leischnig, A., Pflaum, A., Sucky, E. (Eds.), Geschäftsmodelle in der digitalen Welt. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 213–231.

[26] Wrobel, S., Hecker, D., 2018. Fraunhofer-Allianz Big Data, in: Neugebauer, R. (Ed.), Digitalisierung. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 261–273.

[27] Peters, R., 2009. Shopfloor Management: Führen am Ort der Wertschöpfung. LOG_X, Stuttgart, 141 pp.

[28] Hertle, C., Hambach, J., Meißner, A., Rossmann, S., Metternich, J., Rieger, J., 2017. Digitales Shopfloor Management - Neue Impulse für die Verbesserung der Werkstatt. PRODUCTIVITY Management 22 (1), 59–61.

[29] Schmiedgen, P., Tschöpe, S., Nyhuis, P., Noennig, J.R., 2014. Resilienzsteigerung durch Wissenstransfer in CPPS: Dynamischer Wissenstransfer für das Störungsmanagement in Cyber-Physischen Produktionssystemen. wt Werkstattstechnik online 104 (3), 164–168.

[30] Siepmann, D., Graef, N., 2016. Industrie 4.0 – Grundlagen und Gesamtzusammenhang, in: Roth, A. (Ed.), Einführung und Umsetzung von Industrie 4.0. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 17–82.

[31] Ganschar, O., Gerlach, S., Hämmerle, Moritz, Krause, Tobias, Schlund, S., 2013. PRODUKTIONSARBEIT DER ZUKUNFT - INDUSTRIE 4.0. Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO.

[32] Liebrecht, C., Kandler, M., Lang, M., Schaumann, S., Stricker, N., Wuest, T., Lanza, G., 2021. Decision support for the implementation of Industry 4.0 methods: Toolbox, Assessment and Implementation Sequences for Industry 4.0. Journal of Manufacturing Systems 58, 412–430.

[33] Schumacher, A., Erol, S., Sihn, W., 2016. A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises. Procedia CIRP 52, 161–166.

[34] Moosbrugger, H., Kelava, A. (Eds.), 2020. Testtheorie und Fragebogenkonstruktion. Springer Berlin Heidelberg, Berlin, Heidelberg.

[35] Pötters, P., Schindler, P., Leyendecker, B., 2018. Status quo Shopfloor Management. Zeitschrift für wirtschaftlichen Fabrikbetrieb 113 (7-8), 522–525.

Biography

Magnus Kandler studied industrial engineering at the Karlsruhe Institute of Technology (KIT) and has been working as an academic assistant at the Institute of Production Engineering (wbk) at KIT since 2019. His work focuses on the area of production planning and digital Shopfloor Management

Philip Gabriel is currently studying mechanical engineering (B.Sc.) at Karlsruhe Institute of Technology.

Vincent Schröttle is currently studying business engineering (B.Sc.) at Karlsruhe Institute of Technology.

Marvin Carl May studied industrial engineering at the Karlsruhe Institute of Technology (KIT) and has been working as an academic assistant at the Institute of Production Engineering (wbk) at KIT since 2019. His work focuses on the area of production system planning and machine learning.

Prof. Dr.-Ing. Gisela Lanza is member of the management board at the Institute of Production Science (wbk) of the Karlsruhe Institute of Technology (KIT). She heads the Production Systems division dealing with the topics of global production strategies, production system planning, and quality assurance in research and industrial practice.

3rd Conference on Production Systems and Logistics

Agent-based Order Release In Matrix-Structured Assembly Systems

Esben Schukat[1], Jonas Rachner[1], Adrian Maidl[1], Amon Göppert[1], Tobias Adlon[1],
Peter Burggräf[1,3], Robert H. Schmitt[1,2]

[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany*
[2]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*
[3]*Chair of International Production Engineering and Management (IPEM), University of Siegen, Siegen, Germany*

Abstract

The introduction of new variants and the difficulty of forecasting future market demand and developments aggravate the synchronisation of assembly lines. This ultimately leads to cycle time spreads and thus to efficiency losses, e.g. due to lower employee utilisation. In response, matrix-structured assembly systems have been developed as a concept of cycle time independent flow production. Essential characteristics of this type of assembly systems are the dissolution of both one-dimensionally arranged assembly stations as well as cycle times across assembly stations. In recent years, the focus has been on assembly control for the routing of orders through a matrix-structured assembly system. However, order release strategies have largely been neglected, which means that the actually promised performance of this new organisational form of assembly cannot be fulfilled. An agent-based release decision enables the optimal scheduling of new orders taking into account current information from the assembly system such as station states or the processing progress of orders that have already been released. This work extends and builds on existing agent-based approaches to control matrix-structured assembly systems in regard to order release. This results in a theoretical improvement in key performance indicators such as throughput time and station utilisation. For this purpose, the release process, as well as the associated calculation logics and constraints, are described and the implementation in an environmental model is outlined. An essential part of calculation logics is the prediction of all possible paths and capacity requirements resulting from routing and sequence flexibility. This work contributes to the practical realisation and economic operation of matrix-structured assembly systems.

Keywords

Matrix-structured assembly; agile assembly; assembly control; multi-agent system; order release

1. Introduction

Matrix-structured assembly systems are gaining progressively more attention in research. They are considered a potential solution for addressing the ongoing challenges caused by shorter innovation cycles and volatile market conditions through the use of a flexible organizational assembly form [1–4]. Key features of this new organizational form are the breakup of one-dimensionally arranged assembly stations and the elimination of uniform cycle times for all assembly stations [4]. This enables the processing of orders on a situational basis and depending on current circumstances in the assembly system. Resources at assembly stations can not only be used for station-specific assembly steps, but rather for all assembly steps which require this resource. Since the possible routes of an order are known a priori by the assembly system, whereas the actual routing is determined as a response to the situation and depending on the actual availability of the assembly stations, assembly control gains significantly in importance [4–8]. In matrix-

DOI: https://doi.org/10.15488/12139

structured assembly systems, assembly control deals with the assignment of orders and assembly steps to assembly stations, taking into account current resource availability or disruptions. This shifts the complexity from line balancing to assembly control [9,10]. Assembly control deals with the customer- and order-related design of material and information flows. This includes the systematic interaction of material-processing and material-moving areas in a time-related context [11]. Important functions of assembly control are order release and order monitoring, job scheduling, worker assignment, material supply and disruption management [12].

In principle, assembly control can be realized by both centralized and decentralized control architecture. In the application context, centralized architectures are based on extensions of the flexible job shop problem [13–15]. In comparison to decentralized architectures, centralized architectures show better results in simulations with respect to tardiness and lead time [5]. However, for complex problem formulations as well as extensive scenarios, optimal solutions cannot be determined in polynomial time. Centralized control architectures cannot guarantee the requirements of real-time capability [5,8,13–15] which limits their applicability to control matrix-structured assembly systems [5,9]. Accordingly, research in this area has largely focused on decentralized control architectures, which can be well modelled by Multi-Agent systems (MAS) [16–18]. Several studies already have highlighted the efficiency and close to real-time performance of MAS-based assembly control systems [2,5,10,19–21].

Especially in the variant-rich and disruptive operational practice, it is indispensable that orders are only released when the assembly system has foreseeable sufficient capacities for order processing [10]. However, order release as a subtask of assembly control is largely neglected. So far, no functional approaches exist that go beyond random, alternating, or time-based order release. This contradicts the basic principles of matrix-structured assembly systems, according to which orders are only released into the assembly system in case of sufficient resource availability. To realize the performance potentials of matrix-structured assembly systems and to ensure the practicability of agent-based assembly control, a combined view of existing assembly controls and order release is required. Therefore, this paper presents a modular concept for agent-based order release in matrix-structured assembly systems. The concept is derived and described in a system-independent way. It includes the definition of necessary interfaces to the assembly control as well as processes for information management and decision-making in the order release mechanisms. Finally, an outlook on the software implementation is given and the added value and limitations of the concept are discussed.

2. Assembly control in matrix structured assembly systems

2.1 Agent-based control strategies

A MAS consists of several agents that jointly perform one or more tasks through interaction. An agent is a delimitable software unit with defined goals. This unit is embedded in a closed environment and is capable of performing autonomous actions while interacting with other agents in the environment to achieve the defined goals [22]. MAS are particularly characterized by high stability and reliability. They are significantly less likely to fail in dynamic environments than monolithic architectures. Overall, the greater the dynamics and turbulence of the environment, the higher the superiority of MAS becomes [23].

The existing MAS in the area of research at hand differ in terms of their modelling depth and scope, but they can all be considered as possible control systems. Therefore, the approach according to MAYER et al. is exemplarily presented in the following [10]. This approach is based on four agent types: Order release agent, routing agent, workstation agent and vehicle agent. Figure 1 shows these with respect to their interaction and main information flows.

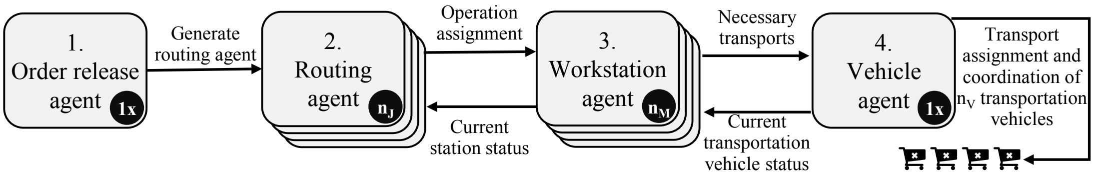


Figure 1: Exemplary decentralized agent-based control approach [10]

The order release agent decides on the timing and product type for the release of orders into the assembly system. The release itself is triggered when the work-in-progress (WIP) falls below a defined threshold. After releasing an order, the routing agent is generated and regularly calculates optimal routing minimizing the individual makespan for the assigned order. The optimization via Monte-Carlo Tree Search (MCTS) is triggered when a routing agent's operation has been finished or by order release. The next type of agent is the workstation agent, which is generated for each workstation in the system and serves to optimize the workstations' schedules and minimizes their idle time between the operations using the same MCTS. The fourth type of agent described by MAYER et al. is the vehicle agent. Like the order release agent, there is only one central vehicle agent to assign and coordinate the transportation vehicles. Its aim is to minimize the total transport costs in terms of time. As an alternative approach, BURGGRÄF et al. distribute the tasks of assembly control among the three agent types: assembly station agent, order agent, and manager agent, the latter also being responsible for order release [5]. However, these alternative approaches differ only in their modelling, while the core functions of assembly control are handled in a similar way.

In the following, the presented and other selected approaches are discussed in the context of the application in order to show concrete deficits in the release mechanisms. However, the basic idea for formulating a release agent according to MAYER et al. seems promising and will be adopted in the following explanations.

2.2 Limitations regarding order release

SCHENK et al. define the verification of resource availability as an essential task of order release [24]. According to GRESCHKE, order release is also an important sub-task of assembly control in matrix-structured assembly systems [8]. It has been shown that MAS are suitable for the control and simulation of matrix-structured assembly systems due to the comparatively faster solution-finding, especially if unexpected disturbances and path deviations are to be expected [2,4,5,8]. Existing approaches for agent-based assembly control neglect feedback from the production operation of the assembly system.

While BURGGRÄF et al. do not specify the order release, SCHÖNEMANN et al. apply a randomized distribution of orders to represent a *worst case* scenario [4,5]. GÖPPERT et. al. also utilize a random distribution approach with a fixed interarrival time to model the order release, similar to that of SCHÖNEMANN et al. [7]. The exemplary presented approach of MAYER et al. also only uses randomized and alternating release and refers to the necessity of advanced order release mechanisms. In the present research field, only one advanced approach could be identified. MÜLLER and SCHMITT provide an approach for sequencing order pools, which is based on the quantification of similarities as well as the minimization of similarities between two successive orders. Accordingly, a static order pool and a fixed order processing sequence are assumed. [25] However, this contradicts the core of the responsiveness of a matrix-structured assembly system just as much as a randomized release of assembly orders without matching capacity supply and demand.

In summary, it can be stated that the complete exploitation of the potentials of matrix-structured assembly requires the development of appropriate approaches for order release. Since the basic idea of MAS is extensibility, a corresponding extension of existing MAS in the context of application is a logical consequence. Therefore, order release tasks must be embedded into existing assembly control systems including the definition of interfaces and release mechanisms. In the following, concrete properties for an agent-based order release are derived and used to propose a general solution.

3. Design of an agent-based order release

3.1 Essential properties

To formulate individual properties of an order release in matrix-structured assembly systems, the deficits of existing control approaches as well as the general performance promises of matrix-structured assembly systems were evaluated using available literature. The findings were then reviewed within the consortium of the AIMFREE research project, which ultimately led to the formulation of six specific properties of a suitable approach. These are presented in the following:

The **first property** is motivated by the assumption of multi-functionality of assembly stations in matrix-structured assembly systems. Multi-functionality describes the ability of assembly stations to perform several different operations along the assembly precedence graph. The resulting increased routing flexibility improves the overall system efficiency [7,26,27]. Thus, capacity calculation must be detailed on an assembly operations level as a single consideration of capacities at the system or station level can lead to bottlenecks in the execution of specific operations. If considered solely at the operation level, assembly stations with comprehensive and multi-functional capability profiles would be scheduled multiple times. Therefore, the capacity situation must be analyzed at both the operation level and the system level. Information at this level of detail must be provided to the order release agent in order to evaluate both levels. The **second property** is the need to consider all possible paths and related capacity demands of an order through the assembly system. First of all, this includes the flexibility of the assembly precedence graph (process sequence flexibility), i.e. the possibilities of processing the operations of an order in different sequences. Furthermore, it includes the possibility resulting from the redundancy of capabilities that the same order selects different paths or stations in a system [15]. In accordance with this flexibility, both unreleased orders and orders that have already been released must be evaluated. In the latter case, all orders in the system must be analyzed in regard to the remaining paths and the capacity utilization to be derived from them. The **third property** is the event-driven release orientation. Periodic release mechanisms are not effective since it has been proven that lead times vary in the application context [28]. The release process needs to be triggered by current events such as the completion of an order. This ensures that current circumstances and information of the assembly system are taken into account for the release decision. The **fourth property** deals with the consideration of individual assembly system operation goals by parameterizing the decision behaviour as proposed by BURGGRÄF et al. [5]. Thereby, due to the agent-based approach of assembly control, the behaviour is encapsulated, but the performance indicators of the assembly system are influenced by order details and released orders. By parameterizing the decision behaviour, individual production strategies can be taken into account, e.g. to minimize lead times, delays or fluctuations in capacity utilization. The **fifth property** is the consideration of order-specific characteristics such as due-dates as well as individual orders [13]. Orders are therefore not grouped in production lots. Instead, in addition to the specification of the product type to be assembled, an order contains further information such as the completion date or margin. This information can also be included in the release decision. The last and **sixth property** is the practicability of the approach regarding solution time [5,9]. An upper boundary can be set by limiting the calculation time of a decision to be significantly shorter than the shortest operation time. Generally, the usage of outdated information needs to be minimized. After formulating comprehensive properties for an agent-based order release, a concrete approach will be presented in the following.

3.2 Embedding of the order release agent

Similar to MAYER et al., the task of order release is embedded in a single order release agent (ORA) and is part of the agent-based assembly control. Together with the matrix-structured assembly system and an order pool the environmental model is formed. The ORA can retrieve or provide information to its surrounding. To initiate the order release process, the ORA can proactively listen to events or be trigged by events in the environmental model. Once triggered, the ORA performs several subtasks related to the decision-making. These subtasks include the request of information such as waiting orders in the order pool. Since the order pool is usually managed by higher planning levels, it cannot be generally interpreted as a direct component of the assembly system or assembly control. Thus, it is located in the environment. However, the order pools provide information to their surrounding. The ORA saves the result of the order release and makes it available to the environment model. The interaction and the functionalities of the ORA are illustrated in Figure 2.

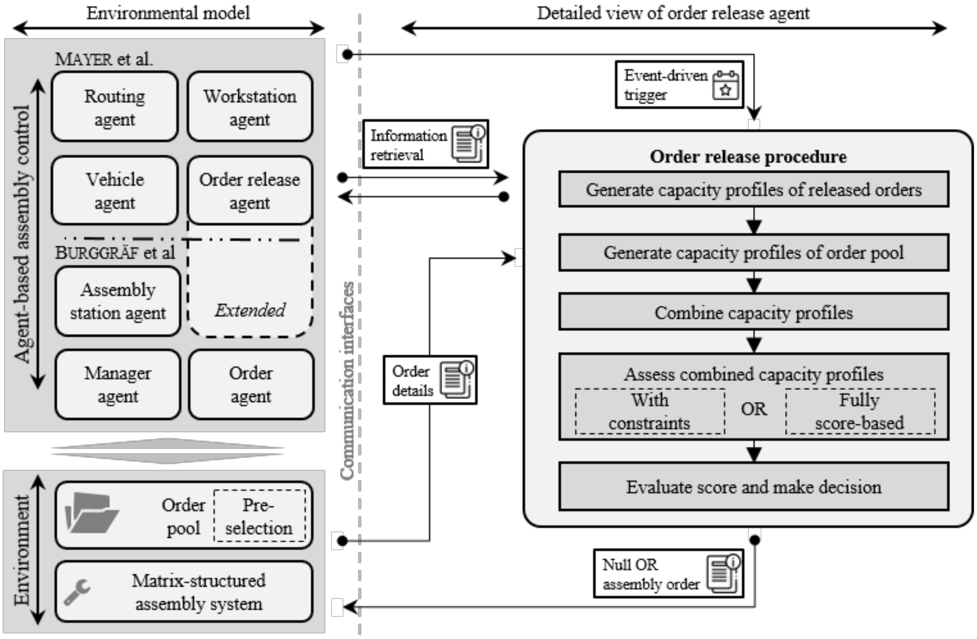


Figure 2: Concept, interactions and functionalities of the order release agent

As soon as the ORA is triggered by a defined event such as order or operation completion, the ORA first wants to replicate the current capacity situation in the assembly system. To do this, it retrieves information about the assembly system from the agent-based assembly control. The information includes available assembly stations as well as their capability profiles. Furthermore, the current order progress of released orders in the assembly system is retrieved. Based on the characteristics of the assembly system, the available capacity can be determined at the operation and system level. Using all path combinations, all the possible capacity profiles of the released orders are generated. In combination with the capacity profiles of the orders in the order pool the order-specific capacity fit can be derived. For this purpose, either an entire order pool or pre-selected groups (e.g. by determining volume cycles [13]) can be used.

The evaluation of the orders with regard to their suitability for release can be done either with restrictions for the capacity limits or through a fully score-based approach. In case of using constraints, overcapacity is not allowed in the system. Consequently, all orders which would exceed the capacity at any given time are rejected. To efficiently verify this, the product specifications of the orders in the order pool are considered first. Then the capacity constraints for all products in the order pool are evaluated and the product with the best capacity fit that does not violate the constraints is selected. In the next step, all orders containing the identified products are filtered. Those filtered orders are score-rated on order-specific information as the due-date and margin. Once all orders are viewed, the best order with the best score is chosen and released. If there is no order which satisfies the constraints or if there is a product which would fit, but no orders with this product exist, no order is released, and the agent goes into standby. Alternatively, a fully score-based approach to capacity evaluation can be used, which neglects hard capacity constraints. The fully score-based evaluation also determines a capacity fit for each product. However, if a product would overfill the available capacity, it’s not set invalid. Instead, a lower capacity fit is assigned to the product. Afterwards, the order pool is filtered for orders containing the product with the best capacity fit. Then, a score based on the weighing of capacity fit and order characteristics is formed. After all orders are rated, the best order is chosen. To avoid that the ORA continuously releases orders, a threshold can be set in the final result-evaluation and decision-making. Through this matching, there is a possibility that no order is released at all. Thus, the fully score-based approach enables the strategic release of orders which could possibly overload the assembly system for a short time by up-scoring certain order details such as due dates. The behaviour of the ORA could be further influenced by adding weighting factors which weight individual operations higher than others, resulting in a better utilization of this operation. Margin, due dates or other order characteristics can be weighted similarly to influence the overall behaviour of the order release agent.

Finally, the results of the capacity evaluation are processed to derive a concrete release decision. If no order was passed down from the capacity evaluation, no release is triggered. Otherwise, a specific assembly order will be released. The release of an order can also be set as a triggering event, so that the described process is initiated again.

3.3 Evaluation of the capacity fit

The main challenge in implementing the ORA and the release mechanisms is the formulation and evaluation of the capacity profiles. In preparation for the presentation of a concrete approach to address this challenge, a fictive case including a nomenclature is illustrated in figure 3.

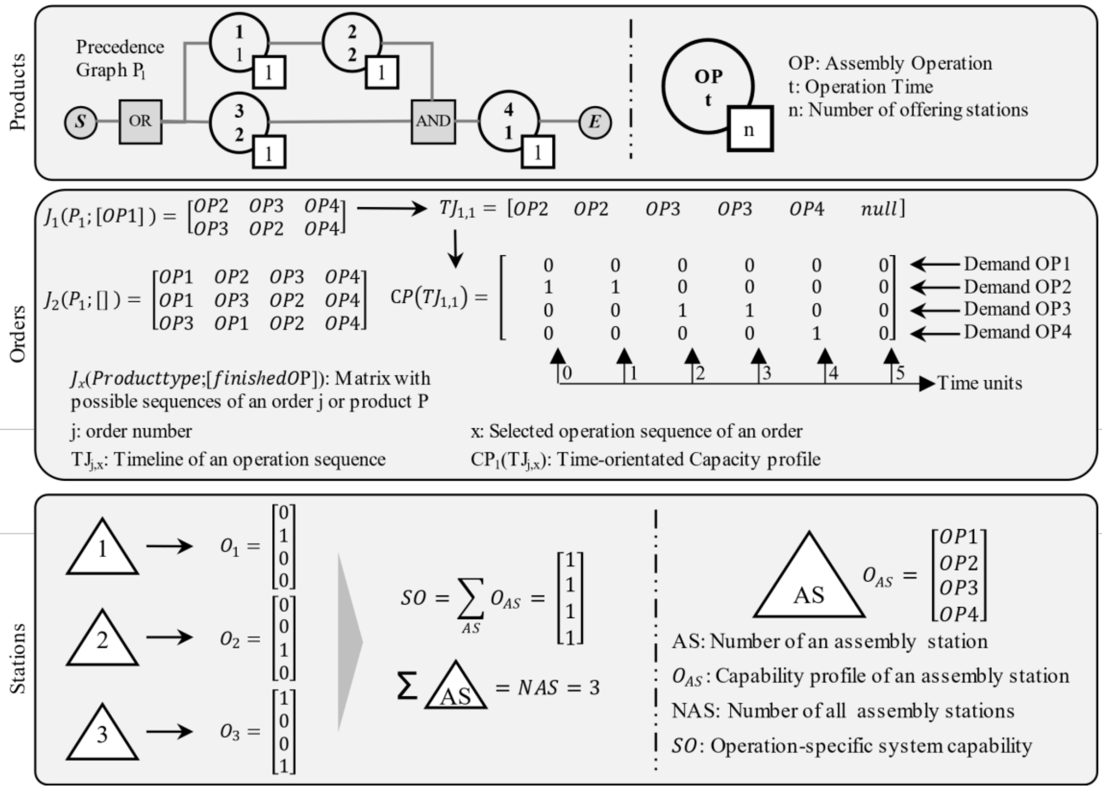


Figure 3: Case scenario and preparation of a capacity profile

Given is a product P1 with four different operations and precedence constraints. For example, the processing time of operation OP4 is one time unit and all other operations must be finished before OP4 can be started. All possible paths can be derived from the assembly precedence graph. These are formulated as a matrix, taking into account operations that have already been carried out, with each line describing a possible processing sequence. The order $J_1(P_1;[OP1])$ has already been processed with respect to OP1, so that two possible processing sequences remain. Now, for example, the first line is used as a possible sequence and gets extended using the given operation durations (see $TJ_{1,1}$). Furthermore, this possible sequence of operations can be transferred into a time-oriented capacity profile by documenting the operation demand line-by-lime (see $CP_1(TJ_{1,1})$). This results in a sequence- and product-specific capacity profile. Furthermore, each assembly station's capability is described with a vector O_{AS} which indicates whether a specific operation can be conducted using a binary variable. For example, the capability profile O_3 permits the processing of OP1 and OP4. The operation-specific system capability is described by the vector sum. Figure 4 illustrates how this formulation of capability profiles is applied to match released orders with new release options.

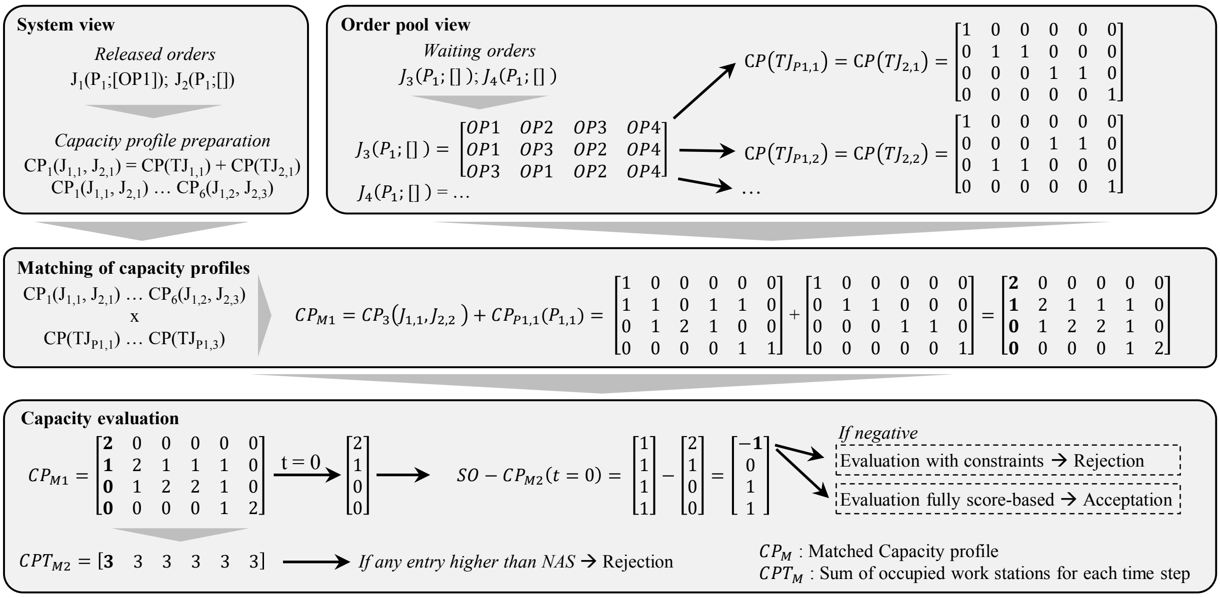


Figure 4: Capacity matching and evaluation

From his surroundings, the ORA finds out that two orders have already been released for processing: $J_1(P_1;[OP1])$ und $J_2(P_1;[])$. As shown in Figure 3, these two orders result in two respectively three possible sequences. Therefore, in a next step, all combinations of the sequences have to be transformed by adding the individual, time-oriented capacity profiles, in this case six profiles. Similarly, from the order pool view, two possible orders are made available for release, in this case each with the product P1. Since no order-specific details (e.g. due dates) are evaluated in the first step, it is sufficient to consider one of the orders on the product level. Just as for the already released orders, the time-oriented capacity profiles are determined. This results in three more profiles, since three possible sequences can be extracted from the precedence graph. The next step is to match the capacity profiles by storing all 18 (six times three) combinations as a capacity profile match. In this way, all possible sequence combinations of the already released orders are matched with those available for release. Then, for a specific time step (here t=0), the demand for the assembly operations is determined and compared with available capacity in the system. In case of a demand overload, depending on the evaluation method (see Figure 2), the matched capacity profile is either rejected or saved as release option with reduced favorability. In order to prevent an overload of the entire system, it is also analyzed whether more stations are occupied by the match than are available. In a similar way, a match is rejected depending on the evaluation method. The process shown must be carried out for all products available in the order pool. Valid matches are prioritized in the following step with regard to further, order-specific characteristics.

4. Discussion

The conception of the ORA bases on an environment model consisting of an agent-based assembly control and order pool. While previous approaches use order lists for random, alternating or lot-based release, the ORA decouples the two aforementioned entities of the environment model. The consideration of current information is essential especially in the context of highly dynamic environments like in matrix-structured assembly systems. Static mechanisms, which release random, alternating, or lot-based orders, are not effective in the intended context. If specific capabilities in the assembly system are overloaded, a lack of consideration of capacities leads to queues, longer throughput times and, in the worst case, deadlocks. At the same time, the use of static release mechanisms can lead to assembly stations in the system remaining completely unused. The extent to which the ORA can access an entire order pool must be decided upon the use case. For this purpose, inventory levels and necessary pre-manufacturing processes must be considered optionally. The danger that individual orders are not drawn from the order pool and remain unprocessed is eliminated by taking individual order characteristics into account and considering them in the decision-making process. The presented approach further requires that all possible order paths are considered for all capacity considerations. However, full path planning creates an NP-hard (NP: non-deterministic polynomial-time) problem, which can have a negative impact on performance in complex scenarios. Thus, approaches for the emerging NP-hardness in path prediction have to be elaborated in order to further enable real-time capable control of matrix-structured assembly systems.

5. Summary and Outlook

This paper has highlighted the relevance of order release to fulfil the performance promise of matrix-structured assembly systems. The presented approach provides a conceptual framework for embedding an order release agent into the agent-based control of a matrix-structured assembly system. However, this paper does not present concrete algorithms. Nevertheless, it can be concluded that in the dynamic environment of a matrix-structured assembly system, random, alternating, or lot-based order release can lead to efficiency losses. Further research should deal with the actual implementation of concrete algorithms as well as the validation of these within practical examples. However, this possibly requires an approach to address the NP-hardness of path planning. A possible solution would be to filter the order-specific paths with regard to improbable path constellations before analyzing them with regard to their capacity profiles. In this way, the capacity analysis could be excluded for rarely occurring paths or, for example, initially very long paths. The path determination is integrated into the ORA in the presented approach. Alternatively, this could also be outsourced to a separate agent in order to generate further performance advantages using asynchronous programming. In addition, a discounting of future capacity loads based on the net present value method should be discussed in order to give more weight to the near and more predictable future when making decisions. The implementation further includes the determination of weighting factors to enable target-oriented and robust decision making. On the basis of the software implementation, the theoretical potential of an agent-based order release can finally be assessed.

Acknowledgements

We would like to thank the BMWi, DLR-PT and our partners for their kind support. This work is part of the research project "AIMFREE", which is funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) within the "Directive on a joint funding initiative to support research and development in the field of electro mobility" (funding code: 01MV19002A) and supported by the project management agency German Aerospace Center (DLR PT). The authors are responsible for the content. More information about the current research activities within the project AIMFREE – Agile assembly of electric vehicles can be found at www.aimfree.wzl.rwth-aachen.de

References

[1] Göppert, A., Schukat, E., Burggräf, P., Schmitt, R.H., 2021. Agile Hybrid Assembly Systems: Bridging the Gap Between Line and Matrix Configurations, in: Weißgraeber, P., Heieck, F., Ackermann, C. (Eds.), Advances in Automotive Production Technology, pp. 3–11.

[2] Greschke, P., Schönemann, M., Thiede, S., Herrmann, C., 2014. Matrix Structures for High Volumes and Flexibility in Production Systems. Procedia CIRP 17, 160–165.

[3] Schmitt, R.H., Göppert, M.R.A., Hüttemann, G., Lettmann, P., Rook-Weiler, K., Schönstein, D., Schreiber, A., Serbest, E., Steffens, M., Tomys-Brummerloh, A., 2017. Frei verkettete wandlungsfähige Montage, in: Brecher, C., Klocke, F., Schmitt, R., Schuh, G. (Eds.), Internet of Production für agile Unternehmen. AWK, 1. Auflage ed. Apprimus Verlag, Aachen.

[4] Schönemann, M., Herrmann, C., Greschke, P., Thiede, S., 2015. Simulation of matrix-structured manufacturing systems. Journal of Manufacturing Systems 37, 104–112.

[5] Burggräf, P., Dannapfel, M., Adlon, T., Kahmann, H., Schukat, E., Holtwiesche, L., 2020. Real-time scheduling of flexible operation sequences on multifunctional assembly stations. wt Werkstattstechnik online 110 (04), 170–176.

[6] Foith-Förster, P., Bauernhansl, T., 2015. Changeable and reconfigurable assembly systems – A structure planning approach in automotive manufacturing, in: Bargende, M., Reuss, H.-C., Wiedemann, J. (Eds.), 15. Internationales Stuttgarter Symposium, Wiesbaden, pp. 1173–1192.

[7] Göppert, A., Rachner, J., Schmitt, R.H., 2020. Automated scenario analysis of reinforcement learning controlled line-less assembly systems. Procedia CIRP 93, 1091–1096.

[8] Greschke, P., 2015. Matrix-Produktion als Konzept einer taktunabhängigen Fließfertigung. Dissertation. Technische Universität Braunschweig.

[9] Burggräf, P., Dannapfel, M., Adlon, T., Riegauf, A., Schukat, E., Schuster, F., 2020. Optimization approach for the combined planning and control of an agile assembly system for electric vehicles, in: Nyhuis, P., Herberger, D., Hübner, M. (Eds.), Proceedings of the Conference on Production Systems and Logistics, Hannover, pp. 137–146.

[10] Mayer, S., Gankin, D., Arnet, C., Endisch, C., 2019 - 2019. Adaptive Production Control with Negotiating Agents in Modular Assembly Systems, in: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy. 06.10.2019 - 09.10.2019, pp. 120–127.

[11] Warnecke, H.J.,, 1986. Ablauf der Montage, in: Spur, G., Stöferle, T. (Eds.), Handbuch Fügen, Handhaben, Montieren, 1. Aufl. ed. Hanser, München, pp. 607–619.

[12] Wiendahl, H.-P., 2014. Betriebsorganisation für Ingenieure, 8th ed. Hanser, München.

[13] Bochmann, L.S., 2018. Entwicklung und Bewertung eines flexiblen und dezentral gesteuerten Fertigungssystems für variantenreiche Produkte. Doctoral Thesis. ETH Zurich.

[14] Buckhorst, A.F., Huettemann, G., Grahn, L., Schmitt, R.H., 2019. Assignment, Sequencing and Location Planning in Line-less Mobile Assembly Systems, in: Schüppstuhl, T., Tracht, K., Roßmann, J. (Eds.), Tagungsband des 4. Kongresses Montage Handhabung Industrieroboter, pp. 227–238.

[15] Burggräf, P., Dannapfel, M., Adlon, T., Schukat, E., Kahmann, H., Holtwiesche, L., 2020. Modeling and evaluating agile assembly systems using mixed-integer linear programming. Procedia CIRP 93, 1073–1078.

[16] Libert, S., Chisu, R., Luft, A., 2010. Softwarearchitektur für eine agentenbasierte Materialflusssteuerung, in: Günthner, W., Hompel, M. ten (Eds.), Internet der Dinge in der Intralogistik, pp. 95–106.

[17] Unland, R., 2015. Industrial Agents, in: , Industrial agents. Emerging applications of software agents in industry, First edition ed. Elsevier, Amsterdam, Netherlands.

[18] Valckenaers, P., van Brussel, H., 2015. Design for the unexpected. From holonic manufacturing systems towards a humane mechatronics society., in: , Industrial agents. Emerging applications of software agents in industry, First edition ed. Elsevier, Amsterdam, Netherlands.

[19] Leusin, M.E., Kück, M., Frazzon, E.M., Maldonado, M.U., Freitag, M., 2018. Potential of a Multi-Agent System Approach for Production Control in Smart Factories. IFAC-PapersOnLine 51 (11), 1459–1464.

[20] May, M.C., Kiefer, L., Kuhnle, A., Stricker, N., Lanza, G., 2021. Decentralized Multi-Agent Production Control through Economic Model Bidding for Matrix Production Systems. Procedia CIRP 96, 3–8.

[21] Tan, Q., Tong, Y., Wu, S., Li, D., 2020. Towards a next-generation production system for industrial robots: A CPS-based hybrid architecture for smart assembly shop floors with closed-loop dynamic cyber physical interactions. Front. Mech. Eng. 15 (1), 1–11.

[22] Wooldridge, M.J., 2009. An introduction to multiagent systems, 2. ed. ed. Wiley, Chichester, West Essex, United Kingdom, 461 pp.

[23] Jeschke, S., 2015. Kybernetik und die Intelligenz verteilter Systeme: Nordrhein-Westfalen auf dem Weg zum digitalen Industrieland, in: Jeschke, S. (Ed.), Exploring Cybernetics. Kybernetik Im Interdisziplinären Diskurs. Springer Fachmedien Wiesbaden GmbH, Wiesbaden.

[24] Schenk, M., 2014. Fabrikplanung und Fabrikbetrieb: Methoden für die wandlungsfähige, vernetzte und ressourceneffiziente Fabrik, 832 pp.

[25] Mueller, D., Schmitt, T.V., 2020. Production planning in autonomous and matrix-structured assembly systems: Effects of similarity of precedence graphs on order release sequencing. Procedia CIRP 93, 1358–1363.

[26] Burggräf, P., Dannapfel, M., Adlon, T., Kahmann, H., Schukat, E., Keens, J., 2020. Capability-based assembly design: An approach for planning an agile assembly system in automotive industry. Procedia CIRP 93, 1206–1211.

[27] Hofmann, C., Brakemeier, N., Krahe, C., Stricker, N., Lanza, G., 2019. The Impact of Routing and Operation Flexibility on the Performance of Matrix Production Compared to a Production Line, in: Schmitt, R., Schuh, G. (Eds.), Advances in Production Research, pp. 155–165.

[28] Küpper, D., Sieben, C., Kuhlmann, K., Ahmad, J., The 2018. Will Flexible-Cell Manufacturing Revolutionize Carmaking? https://www.bcg.com/de-de/publications/2018/flexible-cell-manufacturing-revolutionize-carmaking. Accessed 31 January 2022.

Biography

Esben Schukat (*1993) has been research associate at the Department of Factory Planning at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2019.

Jonas Rachner (*1995) has been research associate at the Department of Model-based Systems at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2019.

Adrian Maidl (*1994) was a master thesis student and student research assistant at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University and has been a Consultant at McKinsey & Company, Inc. since 2022.

Amon Göppert (*1992) has been head of the Department of Model-based Systems at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2021.

Tobias Adlon (*1989) has been head of Department of Factory Planning at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University since 2020.

Prof. Dr.-Ing. Peter Burggräf (*1980) is scientific head of the Department of Factory Planning at the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University and has been head of the Institute of International Production Engineering and Management at the University of Siegen since 2017.

Prof. Dr.-Ing. Robert H. Schmitt (*1961) is head of the Chair of Production Metrology and Quality Management at Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University and member of the Board of Directors of the Fraunhofer Institute for Production Technology (IPT).

3rd Conference on Production Systems and Logistics

MakerSpaces and Value Creation in Start-ups in Germany

Martin Tomczyk[12], Christoph Besenfelder[2]
[1]*Chair for Industrial Information Management, TU Dortmund University, Dortmund, Germany*
[2]*Centre for Entrepreneurship and Transfer, TU Dortmund University, Dortmund, Germany*

Abstract

Initiatives and projects such as the "Excellence Start-up Center.NRW" aim to increase the competitiveness of Germany through startups and has the explicit goal of creating new and sustainable jobs. In addition, so-called MakerSpaces are being created in parallel in many areas, which are considered as creative areas and are intended to support the construction of prototypes and the testing of hypotheses to create value for potential start-ups and to establish a valid business model. The question here is whether these initiatives and projects provide support for industrial value creation in Germany. This would require production and logistics to be considered when creating and developing new business models. Established methods of production research (e.g. simultaneous engineering) and logistics (e.g. supply chain management) should be taken into account. The results of a short survey – by questioning potential startups and advisors – show whether production and logistics are already considered in the consulting by the MakerSpaces or if there are further unmet needs.

Keywords

Value Creation; MakerSpaces; Production and Logistics; Supply Chain Management; Prototyping; Business Models; Survey; Digitalization

1. Introduction

In Germany, the promotion of knowledge-intensive start-ups and innovation culture is seen as a central building block for the further development of the high-tech strategy [1]. The Expert Commission on Research in Innovation (EFI) also recommends expanding and further developing suitable funding formats (EXIST, WIPANO) and promoting appropriate framework conditions, especially for transfer activities from the science sector [2]. The Ministry of Economic Affairs, Innovation, Digitalisation and Energy of North Rhine-Westphalia has responded to this need by funding six "Excellence Start-up Centres.NRW" (ESC) and creating the basis at six universities to develop innovation ecosystems for spin-offs from research. The selected universities are to make a contribution to NRW as a business location with their new support services for start-ups and foundations.

In addition to the role of purely digital-based companies, the preservation and expansion of physical value creation in Europe is seen, especially in Germany, as an important contribution to expanding Germany's international competitiveness and overcoming the multifaceted challenges at a national and global level. In Germany, the political initiative is linked to the successful implementation of a vision of Industry 4.0 [3]. However, the desired industrial revolution is not only associated with major technological changes but always with major social and organizational changes as well [4]. Industry 4.0 is thus not only integrated for the technological advancement of production and supply chain management [5]. A change in organization

DOI: https://doi.org/10.15488/12131

and processes must also be taken into account and thus involve management [6, 7]. In the case of start-ups, the challenges and opportunities of Industry 4.0 must also be considered in the value creation processes and supply chain management when developing new business models and digital processes.

Particularly in start-up funding, the design and production-related issues are addressed in so-called MakerSpaces (MS) and designed for experimentation. Originally from the so-called "Maker Movement" people were led to engage in creative product-making processes by using physical and digital forums to share and collaborate in their daily lives [8]. For the past years, the number of MS increased mainly in the United States and Europe – even worldwide – by more than one thousand[1] [see Fig. 1, 9]. In addition, numerous events such as Hackertons as well as the number of publications increased, especially on the topic of "making" and "hacking" [9].

Therefore, the purpose of this paper is to answer the following research questions:

RQ1: What are the reasons for failure for Teams?

RQ2: Which methods are most commonly perceived by Teams in MakerSpaces?

RQ3: To what extent is production and supply chain management considered in the planning of Teams?

In the following section 2, we discuss the definition of MS as well as the Production requirements for Start-ups. Next, we introduce our short survey and the participants followed by the results in section 4. In Conclusion in section 5, we discuss the results of the short survey.

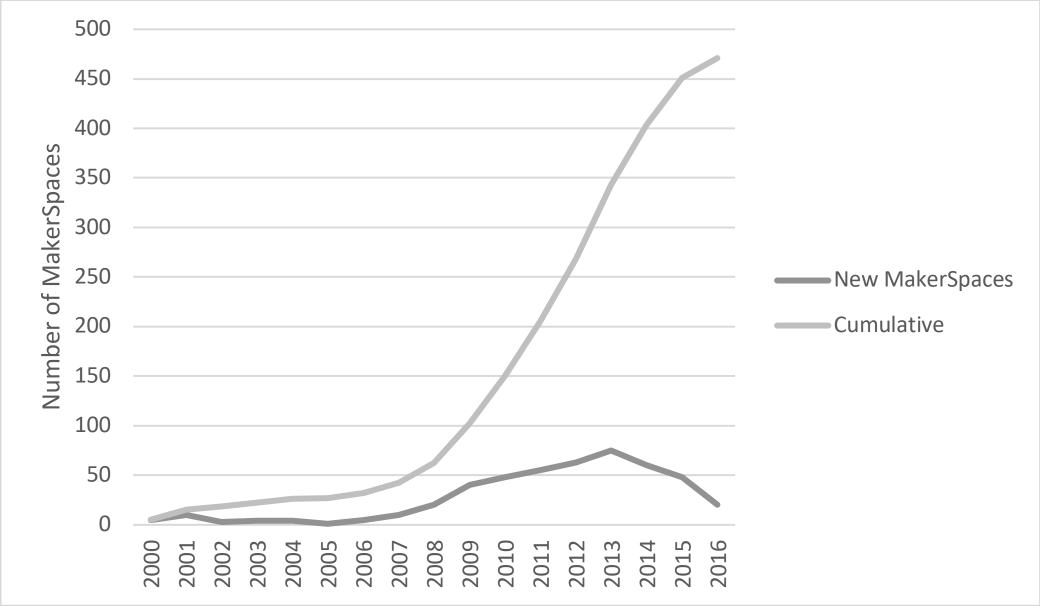


Figure 1: Evolution of the number of MakerSpaces in EU28 per year [9]

---

[1] https://www.popsci.com/rise-makerspace-by-numbers/ - last access January 20th 2022 at 11.09 a.m.

2. State-of-the-Art

2.1 MakerSpaces

One of the first HackerSpaces is considered to be the Homebrew Computer Club 7, which was founded around 1975 in Silicon Valley. Here, hobbyists met informally in a garage to work on do-it-yourself projects, discover technical potential and most likely also discuss politics and society [9]. The term "Maker Movement", on the other hand, ranges from an IT-oriented view to a traditional art-oriented view [9]. This resulted in MakerSpaces (MS), which nowadays have different characteristics as a basis. These can be referred to as FabLabs, HackerSpaces, or MakerSpaces. These characteristics have a spatial offer in common. The basic idea is to offer the community and entrepreneurs an infrastructure so that they can work on everyday problems as well as product ideas and find particularly creative solutions. In addition, they provide a place for free expression, development, and production of expertise and experience. [9, 10, 11]

The focus of MakerSpaces is on publicly accessible creative spaces that put the maker mentality and tinkering practices at the center. The concept refers to any generic space that encourages active participation, collaboration, and thus knowledge sharing among individuals and teams through the original use of technology. In particular, predefined structures are avoided to allow creative processes to be freely designed. [9, 11]

In general, four different goals pursued by an MS can be identified. First, bringing creative technology developments back to communities and cities is an important aspect. The related basic idea of open source and sharing have already led to descriptive innovations in the past. Thus, the proliferation of MS could bring a transformative force. Third, this is accompanied by the vision of creating better integration of science, technology, and business. The idea of spreading craftsmanship early and easily in the community generally leads to increased access to digital manufacturing technologies and tools. Potential startups subsequently support local economies and impact regions positively. [1, 9]

Tailored offerings and collaborative problem-solving in an open-space environment accelerate high-quality manufacturing. Such accelerated prototyping of new, highly customizable products is only possible in MS, in a risk-free and cost-effective manner. Thus, by default, MS are designed to create an environment that encourages the sharing of experiences and expertise. They encourage the use and creation of open content and data, including open hardware and software. Through a creation process based on unrestricted access to documentation, manuals, source code, or design drafts, projects are open to anyone who wants to reuse, revise, remix, and redistribute/process them. [1, 9, 10]

2.2 Production Requirements

Various challenges must be addressed to design an efficient value creation. The design of a business model that takes digitalization and Industry 4.0 into account can be supported by methods such as the Business Model Canvas [12] or Design Thinking [13]. These methods are already used in this context to design the value proposition or the revenue system [14, 15]. However, the resulting adjustments at the level of the design of the value creation processes require deeper methods, comparable to the challenge of existing companies to master the digital transformation [16]. Methods that focus on a systematic analysis of the current state of the company and the required degree of digitalization include the Industry 4.0 Maturity Index [17] or the VDMA Industry 4.0 Toolbox [18]. In the case of a start-up, these frameworks can only provide indications for designing the value creation processes, as they do not yet have any existing processes and need more than this strongly technology-driven perspective. As with existing companies, however, it is unrealistic to assume that a start-up can be built directly only on cyber-physical systems and fully digital solutions [19]. At the same time, the life cycle of a factory far exceeds that of its products, so new planning must at least take current possibilities into account [20].

When building new companies, the same or at least similar questions must therefore be answered as to when digitally transforming existing companies. Especially for later scaling, the planning of the value chain is crucial for cost-efficient product creation. Therefore, supply chain management must also be planned at an early stage.

Supply chain management is defined as "integrated process-oriented planning and control of the flow of materials, information and money along the entire value chain from the end customer to the supplier" [21]. The goals of supply chain management are very relevant, especially when scaling the production of physical goods, as shown by the examples of improved customer orientation, flexibility, reduction of inventory along the value chain, synchronization between supply and demand, and demand-oriented production [21].

To find out how relevant these design fields are for assisted founding teams with physical product ideas and how they are already taken into account in the pre-founding and founding phase, questions were derived based on the tasks of supply chain management.

In general, it was first asked whether the scaling phase had already been considered to assess the maturity of the current planning. The other questions (Are you already in exchange with suppliers?, How is production planned?, Where is production planned?, What are the customers' delivery requirements?) are based on the tactical tasks of the supply chain management task model (network planning). The further questions about inventory planning, supplier selection, and production or assembly capability integrate the operational level of the SCM task model (availability and feasibility check). Finally, the consideration of current challenges regarding sustainability and transparency was queried. The operational level was explicitly neglected since the survey addressed particular start-up projects before series maturity. [21]

3. Survey

The survey is a mixture of qualitative and quantitative research methods. On the one hand, free answers were required, so that subjective opinions of individual persons could be included in the results and the evaluation had to be done in word form. On the other hand, this survey, due to its composition of advisors and startup-oriented teams as well as startups, forms a homogeneous average of relevant and affected persons, so that the survey represents a representative sample. Besides the closed questions and multiple-choice answer options, a clustering of the submitted written answers was initially carried out, which in further steps led to a representation of percent. [22, 23] To be able to cover a basic population of the surveyed participants, an online survey was conducted – also in consideration of the short survey duration (two months). This offered the participants a certain degree of flexibility and allowed the questioner to evaluate the results digitally without delay. [24]

On the one hand, the target groups of the survey were persons and teams interested in founding a company (further referred to as Teams) who are in contact with an ESC (and the Centre of Entrepreneurship and Transfer, short CET) and who use or have used the offers for business model development and prototype development. On the other hand, the target group is composed of supervisors of the first target group as well as employees of the ESC and CET. They are direct supervisors and mentors with access to the industry as well as transfer managers and information and technology managers who have years of expertise in a wide range of industries and fields of activity (further referred to as Supervisors). Thereby a wide spectrum is covered.

In general, some questions were asked to both Teams and Supervisors. On the other hand, due to the different perspectives and backgrounds, some questions were only asked to the individual target groups.

The scope includes more than 40 participants and was conducted from November 2021 to January 2022. Around half of the surveys were not completed. Simply by the fact, that some questions could not be answered by the participants. These incomplete answers were removed from the evaluation. All results come from fully completed surveys (20). Statistically, the number of participants does not indicate any possibility of significance, but in our case, it is negligible due to the participation of experts. Some of the experts have been working together with start-ups and start-up-oriented teams for more than 20 years. This made it possible to capture a broad spectrum and additionally include changes over time.

4. Results

The results of the survey are as diverse as its participants. The remainder of this section will first present the results for each target group and then focus on the results of the common questions.

The results show that more than half of the respondents have a technical-scientific background (e.g., Bachelor/ Master of Science, see Fig. 1). This is due to the proximity of the TU Dortmund University and the adjacent technology center.

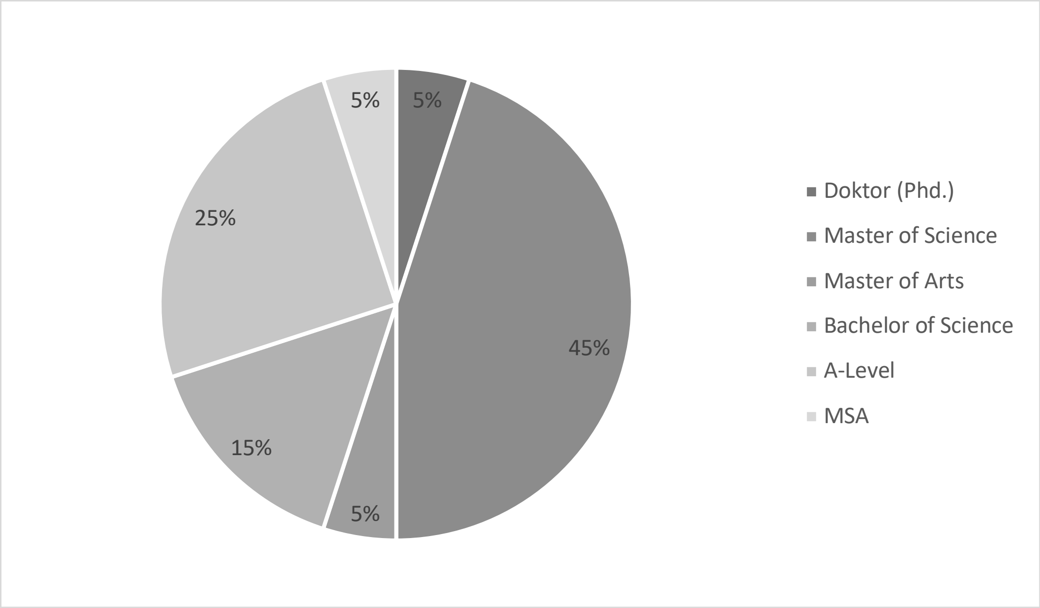


Figure 2: Highest educational degree of participants

On average, in their careers Supervisors have worked or continue to work with 18.93 Teams. In the numbers from supervisors is a wide disparity. The range is from five to 85 Teams. It should be noted that the experts already mentioned above could not specify the number of Teams here. Of the 18.93 Teams, an estimated 1.73 have been founded, although this is based on estimates and the data is not available, particularly for the experts with a high number of Teams under supervision. The planned start-ups of currently mentored Teams are 1.27, so in total there are potentially a total of 3 Teams per Supervisor. In addition, some Teams are in the planning stages of formation. The chances of success among all currently supervised Teams are estimated to be rather average. The range here also extends from one hundred percent conviction to failure. Reasons for this are seen in particular in the product-market fit and the lack of market potential. In addition, a lack of motivation among team members, an incorrect team structure, and a lack of support are given as reasons for the failure of teams (see Fig. 3).

In comparison, the Teams see the greatest problems and challenges in understanding and assessing customer problems and in the lack of innovation and development of the new product. Furthermore, a lack of motivation is also named as one of the main reasons (cf. Fig 4).

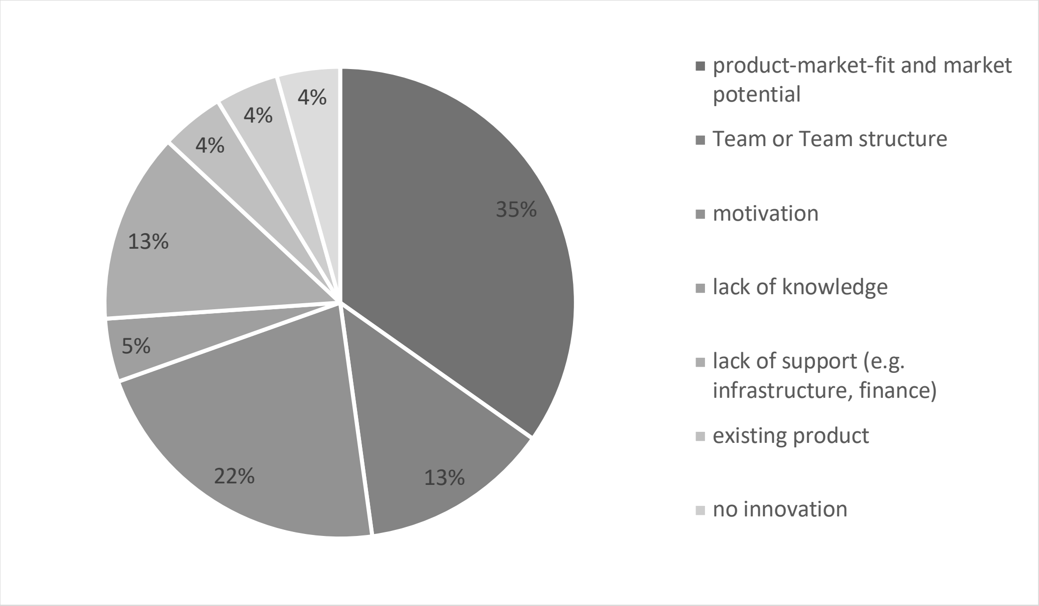


Figure 3: Typical reasons for failure according to supervisors

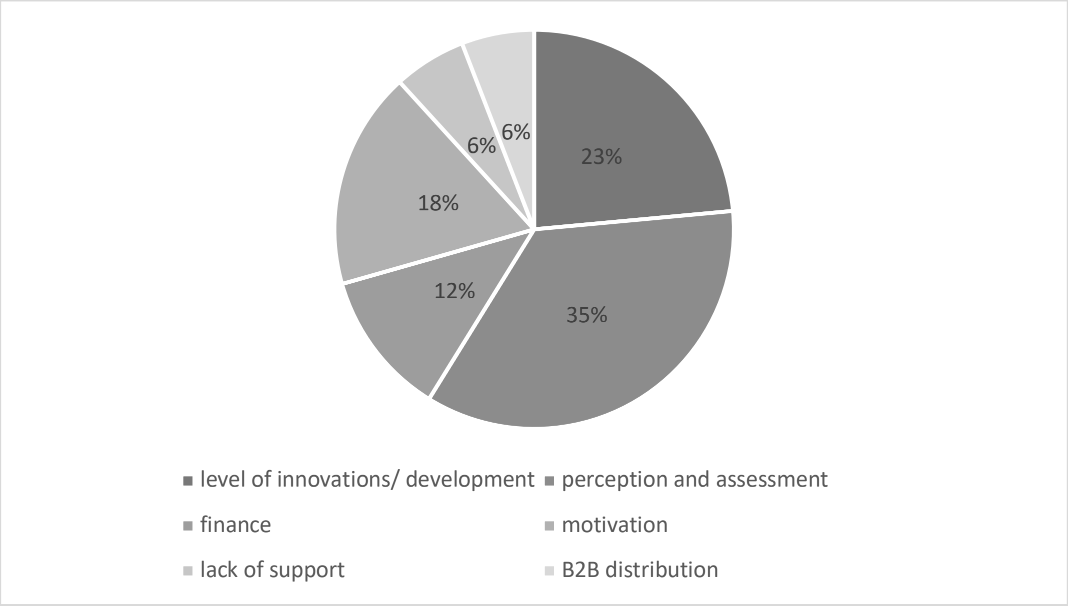


Figure 4: Biggest Problems and challenges according to the Teams

The current financing of the Teams is mostly granted through a funding scholarship. This is primarily the NRW start-up scholarship. Other options are the “EXIST Gründerstipendium” or a validation program such as VIP+, which are accompanied by the future possibility of an investor entry. In addition, fully self-financing takes place in 10% of the cases, which has the disadvantage that the focus is not completely placed on the foundation. In addition, a mixture of both options can take place.

In general, the teams consist of 3.6 members. In this survey, there was no founding of individuals.

Regarding business planning and business model development, the Business Model Canvas is mainly used. Since it was developed for business model development, this is not surprising. Other approaches are the Lean Canvas or the "jobs to be done" method. Design Thinking is primarily used to identify customer needs. Alternatives for this are stakeholder analysis or hypothesis testing. Details are always neglected in the development of the products and the business model. In particular, the Teams stated that they neglected customer interests, their business financing, or the technical realization. Other areas are marketing or product extensions for the future (see Fig. 5).

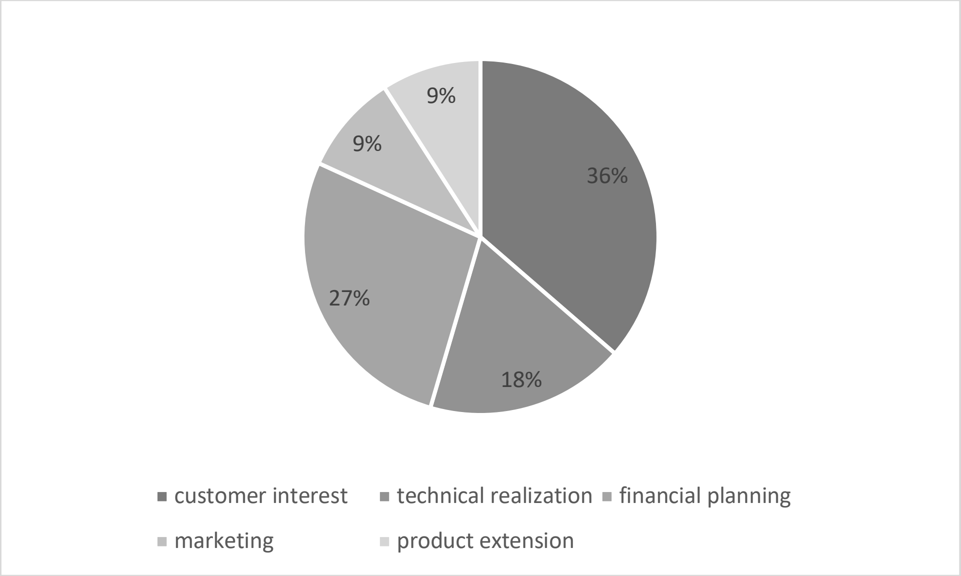


Figure 5: Deliberately neglected according to the Teams

In most cases, the production of a prototype is driven forward in parallel with the development of the business model. Only a quarter of the teams deal with a "make-or-buy" decision. Prototyping is mainly done in digital form (as the development of Applications or via AutoCAD/ 3D model) or using 3D printing. Consideration of future production and delivery is affirmed by 57% of the Teams. Only 43% do not consider this initially (see Fig. 6). Furthermore, few Teams take supplier information into account during their planning. Indeed, five out of six are planning their production within Germany.

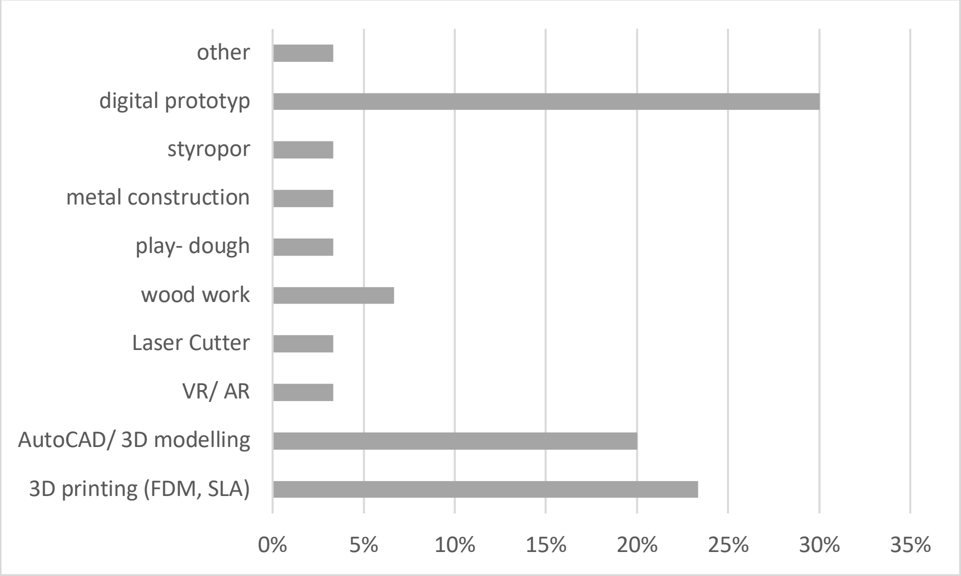


Figure 6: Methods used for prototyping

5. Discussion and Limitations

First of all, it is striking that the chances of success are rated moderately well by the supervisors. On a scale of 1 (very good) to 6 (no chance), the average is 3.06. The reasons here may lie in differences in perception. Many Teams are convinced of their idea and reject criticism as "unfounded" or “unjustified”. Furthermore, the Teams believe they understand the customer problem. Interestingly, the lack of market fit as well as the market potential is mentioned as the main reason for the failure by supervisors and equally listed as a problem and the biggest challenge by the Teams. Additionally, Teams report intentionally not addressing customer needs and technical implementation at the beginning. In this regard, new business models cannot succeed in reality without incorporating customer needs and the appropriate technical implementation. Reasons for this approach of the Teams are explained by the supervisors with a wrong self-perception, a certain naivety, and a lack of trust in the supervisors. Furthermore, the Teams produce their prototypes with the help of the MS. The Teams mainly (50%) use digital services, which can be explained by the increasing number of digital solutions and internet-based start-ups [25]. Further the results show that most of the start-ups with physical product ideas indicate that they are already thinking about the scaling phase, but hardly consider the tasks of supply chain management systematically. It should be emphasized that the majority are planning production in Germany and therefore the assumption of a positive influence of start-ups on industrial value creation in Europe is plausible.

Success factors can be derived as a reverse conclusion. Teams should therefore be realistic, rely on their mentors, and have a clearer focus on customer needs. In addition, production planning is an important element to take into account efficient production and later scaling.

Why the design fields of supply chain management are not yet systematically considered in the pre-start-up phase should be the subject of further research. Since the Design Thinking and Business Model Canvas methods are used very frequently, the conclusion is obvious that a supplementary methodology to highlight the design fields of supply chain management for start-up projects could make a contribution.

In summary, the Teams mostly need their support in identifying customer problems and maintaining motivation. Furthermore, it has been noticed that the number of digital solutions and digital business models is increasing. Therefore, an expansion of MS towards a combination with DataSpaces should be considered. A larger offer for a low-code environment as well as the support regarding digital and internet-based solutions should be investigated in further research work.

Acknowledgments

Excellence Start-up Center. Our project is supported by the state government: At the beginning of 2019, the NRW Ministry of Economic Affairs designated the CET as an Excellence Start-up Center (ESC).

References

[1] Hightech-Forum (Hrsg.) (2021): zusammen. wachsen. gestalten. Ergebnisbericht des Hightech-Forums. Empfehlungen zur Weiterentwicklung der Hightech-Strategie 2025, Berlin.

[2] EFI – Expertenkommission Forschung und Innovation (2021): Gutachten zu Forschung, Innovation und technologischer Leistungsfähigkeit Deutschlands 2021, Berlin: EFI.

[3] Kagermann, H., Wahlster, W., Helbig, J., 2013. Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0. Abschlussbericht des Arbeitskreises Industrie 4.0, o. V., Frankfurt a. M.

[4] Hirsch-Kreinsen, H., 2014. Wandel von Produktionsarbeit - Industrie 4.0, in: WSI-Mitteilungen, 67 (2014) 6, pp. 421–429.

[5] Henke, M., Besenfelder, C., Kaczmarek, S., Fiolka, M., 2020. A Vision of Digitalization in Supply Chain Management and Logistics. Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020).

[6] Henke, M., Besenfelder, C., Kaczmarek, S.: Dortmunder Management-Modell, in: ten Hompel, M., Vogel-Heuser, B., Bauernhansl, T. (Eds.): Handbuch Industrie 4.0, Bd. 3: Logistik, 3. Aufl., Springer Vieweg, Berlin, Heidelberg, 2020, p. 555-571.

[7] Henke, M., Hegmanns, T., 2017. Geschäftsmodelle für die Logistik 4.0: Herausforderungen und Handlungsfelder einer grundlegenden Transformation, in: Birgit Vogel-Heuser, Thomas Bauernhansl und Michael ten Hompel (Hg.): Handbuch Industrie 4.0 Bd.3., Springer, Berlin, Heidelberg, S 335–345

[8] Halverson, E.R., Sheridan, K.M., 2014. The Maker Movement in Education. Harvard Educational Review 84 (4), 495-504.

[9] Rosa, P., Ferretti, F., Perreira, Â.G., Panella, F., Wanner, M., 2017. Overview of the Maker Movement in the European Union, Publications Office of the European Union, Luxembourg.

[10] Hatch, M. (2014). The maker movement manifesto: Rules for innovation in the new world of crafters, hackers, and tinkerers. McGraw Hill Professional.

[11] Anderson, C. (2013). Makers: das Internet der Dinge: die nächste industrielle Revolution. Carl Hanser Verlag GmbH Co KG.

[12] Osterwalder, A., Pigneur, Y., 2010. Business Model Generation. A Handbook for Visionaries, Game Changers, and Challengers. Hoboken, NJ: John Wiley & Sons.

[13] Garbuio, M., Lovallo, D., 2018. Design Thinking, in: Mie Augier und David J. Teece (Hg.): The Palgrave Encyclopedia of Strategic Management, Bd. 8. London: Palgrave Macmillan UK, pp. 1–2.

[14] Schuh, G., Riesener, M., Prote, J. P., Dölle, C., Molitor, M., Schloesser, S., ... & Tittel, J. (2020). Industrie 4.0: Agile Entwicklung und Produktion im Internet of Production. In Handbuch Industrie 4.0: Recht, Technik, Gesellschaft (pp. 467-488). Springer, Berlin, Heidelberg.

[15] Reischauer G., Schober L. (2016) Industrie 4.0 durch strategische Organisationsgestaltung managen. In: Obermaier R. (eds) Industrie 4.0 als unternehmerische Gestaltungsaufgabe. Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-08165-2_16

[16] Henke, M., Parlings, M., Besenfelder, C., Stute, M. & Brandl, T.: Management der Industrie 4.0. (VWI Fokusthema-Band 1). Berlin: Verband Deutscher Wirtschaftsingenieure e.V. 2019.

[17] Schuh, G., Anderl, R., Gausemeier, J., ten Hompel, M., Wahlster, W., 2017. Industrie 4.0 Maturity Index. Die digitale Transformation von Unternehmen gestalten. Acatech Studie. München.

[18] Anderl, R., 2015. Leitfaden Industrie 4.0. Orientierungshilfe zur Einführung in den Mittelstand. Frankfurt am Main: VDMA-Verl.

[19] Henke, M., 2016. 10 Thesen zum Management der Industrie 4.0. Hg. v. Wirtschafts Woche.

[20] Müller, E., Engelmann, J., Löffler, T., Strauch, J., 2009. Energieeffiziente Fabriken planen und betreiben. Springer, Berlin, Heidelberg.

[21] Kuhn, A., Hellingrath, B., 2002. Supply Chain Management. Optimierte Zusammenarbeit in der Wertschöpfungskette. Springer, Berlin, Heidelberg.

[22] Möhring, W., Schlütz, D., 2012. Die Befragung in der Medien- und Kommunikationswissenschaft – Eine praxisorientierte Einführung. 3. Aufl. Springer, Leipzip.

[23] Krosnick, J.A., Presser, S., 2009. Question and Questionnaire Design, in: Handbook of Survey Research (2nd Edition) James D. Wright and Peter V. Marsden (Eds). San Diego, CA: Elsevier.

[24] Evans, J.R. and Mathur, A. (2005), "The value of online surveys", Internet Research, Vol. 15 No. 2, pp. 195-219. https://doi.org/10.1108/10662240510590360

[25] Metzger, G., 2020. KfW-Gründungsmonitor 2020. KfW Research. Frankfurt am Main.

Biography

Martin Tomczyk, M. Sc. (*1993) is a researcher and PhD student at the Chair for Industrial Information Management (IIM) at TU Dortmund University. He is also currently working as head of Data Space at the Centre for Entrepreneurship and Transfer (CET) and has graduated with a master's degree in Industrial Engineering at TU Dortmund University.

Dr.-Ing. Christoph Besenfelder (*1984) studied logistics engineering at TU Dortmund University and subsequently worked as a research assistant at the Chair of Enterprise Logistics (LFO). After completing his doctorate in engineering, he has been research manager at the Fraunhofer Institute for Material Flow and Logistics (IML) for the Enterprise Logistics Division. He is Head of the Innovation and Transfer Management Team at the Centre for Entrepreneurship and Transfer at TU Dortmund University since 2020.

3rd Conference on Production Systems and Logistics

A Holistic View On Production Systems Management

Mark Grothkopp[1], Matteo Bernasconi[1], Thomas Friedli[1]
[1]Institute of Technology Management, University of St.Gallen, St.Gallen, Switzerland

Abstract

Almost every large corporation nowadays operates some sort of Production System (PS), usually built as a derivative from leading examples like the Toyota Production System. Production Systems (PSs) are introduced to increase operational performance and to eventually instill a culture of continuous improvement across the mostly globally dispersed production networks. The main question is not any longer if PSs are helpful but how to manage them. So far there is neither an answer to this question in practice nor in literature. That is, how to design and develop the *content*, the *process* and the organizational support *structure* of a PS, and thus providing a corporate perspective to managing PSs, is heavily under researched.

The methodological approach in this paper is twofold. First, a systematic literature review was conducted to identify appropriate papers dealing with this topic. Second, we draw on interviews with corporate representatives being accountable for the PS at 11 respective companies from the Pharmaceutical industry. The companies have been selected based on their maturity of production system implementation. Interviews were transcribed and coded.

We found various activities related to the three dimensions of *content*, *process*, and *structure* of PSs. Thereby, we provide an overview of activities for managing PSs. We add to the literature of PSs from a corporate perspective and derive several future research opportunities, such as if there are multiple ways in combining the identified activities to be successful with a PS. Our limitation is that interviewees are from the pharmaceutical industry only, yet the level of sophistication of PSs in this industry and the twofold approach mitigate the limitation.

Keywords

Production Systems; Continuous Improvement; Corporate; Lean; Pharma Industry

1. Introduction

Since the industrial revolution, manufacturing companies have deployed programs to improve their operations. The pioneering approach of Toyota and the creation of the world-famous Toyota Production System set the basis for lean production [1]. Automobile manufacturers started copying and developing their company specific Production Systems (PSs). Over time, PSs have become popular beyond the automotive industry and manufacturing departments [2]. The main goal of deploying PSs remains the same across all industries. Companies aim at improving effectiveness and efficiency in their operations and developing a culture of continuous improvement [3].

However, reaping the promises is not as easy as it sounds [4]. Several companies eventually fail with their Production System (PS) [5–7]. At others, the PS is only superficially integrated and/or only seen as a "toolbox", and hence limited in driving continuous improvement systematically [8,9]. Hence, companies are faced with the question of how to manage their PS [3].

DOI: https://doi.org/10.15488/12161

This paper aims at examining *what are activities to manage a PS*. Firstly, we performed a systematic literature analysis to identify the relevant papers dealing with Production System management. Secondly, we conducted interviews with representatives from 11 pharmaceutical companies. Finally, we present our findings and show future research potentials.

2. State of Research

Many companies have adopted or even copied some parts of the Toyota Production System [10,11] and adjusted the design and structure [12], thereby generating their own interpretation of a PS [13]. There is no coherent definition of a PS, a PS is e.g. seen as a framework consisting of strategic goals, principles, and tools to manage production processes [14,15] or as an improvement program [16]. Since programs need to be managed [17] and this article aims at identifying activities to manage PSs, the definition used by Netland [16] and subsequently other researchers [18,19] of an "improvement program" is being followed in this article. Additionally, this article takes a perspective from the headquarter as PSs are often initiated by the headquarter and deployed to sites [20,21].

By creating their PS, companies aim at a structured implementation of common practices and the creation of a culture of continuous improvement within their manufacturing network [3,20]. In case of a successful implementation, companies will benefit from improved operational performance [22], achieve a continuous improvement capability [21], and therefore experience a competitive advantage [11]. Although companies use different names to refer to their own PS, they all are in essence a "*process improvement program*" [16]. Netland recognizes three dimensions that describe a PS [16]: the *content* comprises the principles included in the PS and clearly stated in the respective visualization model. The *process* dimension summarizes the mechanism how the PS and its content is implemented. Finally, the *structure* defines the dedicated teams on corporate and site level that are needed to support the process and content dimensions [16]. Literature provides little guidance on how to manage a PS. Academics [2,10,23] point out that companies must align strategy and *content* of PSs, as well as integrate isolated initiatives into the *content* of the PS. In the *process* dimension, Saunders et al. highlight the importance of soft management skills to deploy strategic initiatives [24]. Hekneby et al. focus on the role of managerial attention to sustain the global adoption of PSs [18]. Stalberg and Funding perform a case study with a company and derive several managerial implications for the *process* and *content* dimensions, ignoring the *structure* dimension [23]. Still only few publications discuss a more holistic approach to manage PSs. Albeit Netland [16] specified guidelines for action for a PS in the dimensions *content*, *structure* and *process*, they are neither systematically derived nor collectively exhaustive.

3. Methodology

The research question is answered in a methodological twofold approach, combining a theoretical and practical perspective. The theoretical perspective is based on a systematic literature review. Due to the topic of this research, the systematic literature review follows the step-by-step approach for operations management by Thomé et al. [25]. Two databases namely *ScienceDirect* and *Web of Science* were used for the search which was carried out in June and July 2021. Search words[1] were carefully selected and discussed among the authors. Only peer-reviewed journal articles were included to ensure high-quality results. Papers should cover entire improvement programs, not isolated or single practices, at manufacturing companies. Both forward and backward searches were used to refine the final selection of papers. A total of 39 papers

---

[1] ("OPEX program*" OR "improvement program*"OR "CI program*" OR "JIT program*" OR "TQM program*" OR "TPM program*" OR "Lean Production System*" OR "Toyota Production System" OR "Company-specific Production System*") AND ("sustain*" OR "implement*" OR "manag*")

has been identified during the systematic literature review. Data gathering from the papers was conducted using a concept matrix according to Webster and Watson [26] in order to find activities in managing PSs.

The practical perspective uses semi-structured case interviews which provide the researcher with the opportunity to have a structure in place but also allows for enough flexibility to cope with the complexity of the topic at hand. The selection of the case interview partners was based on three criteria to guarantee the reliability of the process. First, prior engagement between the companies and the authors had to have taken place. This ensures common understanding of a PS, reducing the risk of misunderstandings [27]. Second, partners should be active in pharmaceutical companies. Even though the pharmaceutical industry was a laggard in introducing PSs compared to other industries, it is now at the forefront with regard to the applied sophistication in designing PSs [28,29]. Third, cases should embrace companies with different sizes and maturity stages to mitigate the risk of a selection bias. The semi-structured interviews were performed in 2021. The process relies on a pre-defined interview guideline and questions emerging during the interviews to increase the reliability of the research [30]. This guideline was structured according to Netland's [16] *content*, *process* and *structure* dimension of PSs and pre-tested with a practitioner from one of the case firms. A total of 12 interviews from 11 companies[2] with interviewees from corporate teams were conducted, ranging from 37 to 69 minutes. The size of companies ranges from 2000 to 95000 employees. Interviews were transcribed using the commercial software *trint* and codified with the software *Atlas.ti*. Data coding was structured in two cycles, according to Saldaña [31]. Initially, codes are found using the open-coding technique, which is suitable due to the explorative, open nature of the research [32]. Successively, the initial codes are aggregated in themes by removing redundancies and categorizing them with the axial coding technique [31]. Both methodological approaches were then combined afterwards to refine and structure the selection of activities and sub-activities for managing PSs. The overarching structure comprises *content*, *process*, and *structure* dimensions according to Netland [16].

4. Results & Discussion

4.1 Content

Table 1 contains the activities identified in the *content* dimension.

Table 1: Collection of activities for managing the content of PSs

Activity	Sub-Activity	Literature	Interviews
PS Content Identification	Copying PS Content from External	[33,34]	[A, F, J, K]
	Integration of Existing Projects & Initiatives	[35–37,23]	[A, B, C, E, F, G, H]
	Strategic Alignment of PS & Company Strategy	[5,33,38,16,11,39,40]	[A, B, C, E, F, G, H, I, J, K]
PS Content Definition	Definition of Lean & Fundamentals Elements	[41]	[C, D]
	Combination of Soft & Technical Elements	[10]	[D, E, G, H]
	Integration of Digitalization Elements	[42]	[D, I, J]
	Operationalization of PS Elements	[43,44]	[C, H]
PS Integration	PS Integration with other Improvement Programs	[45,10,46]	[B, D, E, F]
	PS Definition as an Umbrella System	[36]	[A, B, C]

---

[2] Companies 1-11 are referred to as A to K respectively in the following chapters

PS Content Evolution	PS Scope Expansion	[47,39]	[A, B, C, I, K]
	PS Content Adaptation	[48,47,16,39,23]	[A, C, D, E, I]
	Adaptation and Update of Guidelines/Standards	[16]	[A, E, H, I, J]
Adaptation of PS at Sites	Adaptation of Guidelines/Standards at Sites	[44,49]	[A, D, E, G, H, I, J]
	Adaptation of PS itself at Sites		[C, G, K]

We found that companies are not only concerned with how they design their PS initially but also how they adapt it over time or to their sites in the network. The identification of PS elements can be driven by copying from other renowned models, building upon existing improvement projects or initiatives, and by an alignment whit corporate strategy. The integration of PSs relates to how the PS and other on-going initiatives in companies compete for resources and attention from management, thus taking an outside-in perspective from the PS. Another aspect is the PS as an umbrella system enveloping all improvement initiatives, thereby fostering the internal integration of the PS. During the definition of the PS content, that is the description and operationalization of elements, companies shape their PS. From both the practical and theoretical perspective, the content itself comprises lean and other fundamental elements but also contains softer as well as digitalization elements. The operationalization of the PS content refers to the level of detail, ranging from generic to specific. Despite the definition of the content at a company's headquarter, the PS itself and its elements in form of operationalized guidelines or "playbooks" is adapted to the company's sites. While literature shows evidence for the adaptation of guidelines, since they provide a certain degree of freedom to adhere to, it does not for the adaptation of the entire system, such that different elements are merely applicable or applied to selected sites due to their characteristics. The interviews in this case revealed that some companies design their system in a modular way in which some elements are deployed to sites with certain roles. Lastly, we identified the evolution of the PS content which is the adaptation of the PS system itself, the expansion of the scope, and the adaptation of guidelines and playbooks.

4.2 Process

Table 2 includes activities and sub-activities from the process *dimension* of PSs.

Table 2: Collection of activities for managing the process of PSs

Activity	Sub-Activity	Literature	Interviews
Roll-Out Strategy Definition	Roll-Out at Organizational Levels	[47,46]	[A, B, F, H, J]
	Roll-Out Responsibility Definition		[A, B, C, D, E, F, G, H, K]
	Roll-Out Strategy Definition for Network		[A, E, G, I, K]
	Roll-Out Strategy Definition at Sites	[50,36,51,27,52]	[A, C, D, E, F, G, H, I, J, K]
Capability Building & Transfer Mechanisms	Coaching	[53]	[B, D, I, J]
	Training (Workshops)	[50,43,33,44,36,53,54,37,27,52,40,55]	[A, B, D, E, F, J, K]
	Job Rotation	[50,43,36,7,40]	[D, I]
	Setting up Learning Platforms		[D]
	Successful Practice Sharing	[50,43,44,36,54,52,39]	[B, F, H, K]
	Establishing Teamwork (cross-functional)	[50,43,36,56,40,55]	[B, E]

	Definition of Standards, Guidelines & Playbooks	[44,36]	[B, C, E, F, H, J]
Monitoring & Controlling	Ensuring Resource Availability & Financial Support	[51,57,37,52,41,40,55]	[G]
	Performance Management	[50,44,36,51,53,37,41,58,39,40,55]	[A, C]
	Maturity Assessment	[50,43,44,36,53,27,52,58,39,55]	[A, C, D, E, F, G, H, I, J, K]
	Strategy Deployment	[36,51,56,53,59,39,7,40]	[A, B, C, E, F, G, H, I, K]
	Establishing Program & Project Management	[40,55]	[B]
	Daily Visual Management (Huddles)	[53,58,7]	[E, J]
	Conducting Follow-Ups	[35,51,52,55]	[A, I]
	Conducting Gemba Walks	[16]	[D, G]
Management Engagement	Getting Site Leadership Team Commitment	[27]	[A, I]
	Getting Top Management Commitment	[60,37,47,61,16,52,41,40,55]	[A, C, E, F, G, I, J, K]
	Establishing Management Involvement	[51,56,47,52,46]	[A]
	Creating Management Push	[5,36,51]	[A, D, E, J]
Buy-In & Motivation Creation	Creation of a Joint Vision	[44,37,23]	[A, E, I, J]
	Common Language Definition	[36,54]	[A, E, J]
	Raising Awareness	[48,53,54,47,52,39,40]	[B, F, G, J]
	Communication	[50,62,51,37,52,40]	[A, E, F]
	Quick Improvement Benefits Sharing	[5,60,27,11]	[B, C, E, F, G, H, J]

Five activities were identified in the process dimension. First, companies define more or less intentionally their roll-out strategy. Interviews revealed that this design of the roll-out strategy is manifold, covering not only the roll-out within sites but also across the network and organizational levels. That is, whether all sites are in scope at the same time or some might act as a starting point. Companies might also start from top management downwards or start from the bottom upwards with their roll-out. Second, the knowledge or the content of the PS needs to be transferred to the sites and capabilities need to be built at site level. In this activity, various transfer mechanisms are leveraged by companies. Usual mechanisms, such as trainings as well as guidelines and manuals, are used frequently but also other more advanced ones, such as coaching and successful practice sharing, are utilized. Interviews revealed that companies also leverage dedicated learning platforms or online academies. In some cases, a formalized job rotation program was used to systematically build capabilities at the sites. Third, the progress and effectiveness of the PS is monitored and controlled. Most of the sub-activities are conducted in a continuous manner. Ensuring sufficient resources in form of financials and time is critical. Strategy deployment includes goal setting and cascading so that objectives are clear but also well aligned. In addition, maturity assessments are regularly conducted to track the progress of PS implementation. Also, the achieved improvements and potential gaps to be closed are monitored with performance management. Daily management in form of huddles, regular follow-ups and gemba walks are used for controlling and monitoring. Fourth, creating engagement from management was found to be decisive in both literature and practice. Especially, a dedicated push from top management at the beginning of the PS was mentioned in the interviews. Yet, a continued commitment and eventually a revitalizing push from management was also observed in both interviews and literature. Fifth, companies need to create buy-in and motivation within the organization. Frequent communication, raising awareness and a joint vision help to get employees on board for the PS. Additionally, a common language ensures the same understanding and thus further helps in engaging employees.

4.3 Structure

Table 3 summarizes the activities for managing the structure of the PS. The structure of the PS can be divided into formal and informal structures as well as the respective responsibilities. For the formal structure, a corporate PS team is usually established to initiate the PS such as designing its content while site PS teams are leveraged to deploy it. Interviews with practice additionally showed the importance of the team leader of the corporate PS team who is appointed at the beginning of the PS.

Table 3: Collection of activities for managing the structure of PSs

Activity	Sub-Activity	Literature	Interviews
Formal Structures Definition	Central/Corporate Team	[43,36,16,40]	[A, B, C, D, E; G, I, J, K]
	Site PS Teams	[52,49]	[A, E, K]
	Head of Corporate PS Team		[I]
	External Support (Consultancies)	[51,52,41]	[A, E, G, H, J]
Informal Structures Definition	PS Champions/Ambassadors	[33,44,36,51]	[A]
	Communities of Practices		[F]
Responsibilities & Relations Definition	Connection between Corporate & Site PS Teams		[J]
	Ownership & Responsibility of Improvements	[63,57,45,16,7,55]	[D, E, G, I, J, K]

In addition to formal structures, companies establish informal structures or let them emerge organically. These informal structures are PS champions, which are assigned facilitators of the PS but not dedicated to corporate or site PS teams, and communities of practices, which often reflect a group of people interested in a certain topic of the PS. Lastly, the definition of responsibilities for improvements and ownership is a critical activity which is an on-going task. Yet where the responsibility resides might change over time. Moreover, the connection or responsibility between corporate and site PS teams was highlighted by practitioners.

5. Conclusion and Future Research

Companies have experienced improvements with the help of PSs but some of them failed eventually. The question that emerged and that is being raised by companies is how do PSs need to be managed. Our article sheds light on this question by providing an overview of activities for managing PSs. Our twofold approach revealed various activities and sub-activities that companies can consider to manage their PS. These include how they define and adapt the PS content, how they monitor and control the progress, how they build the capabilities at their sites, how they create management engagement and organizational buy-in, and how they establish the respective organizational structure. With this, we add to the literature of PSs by integrating single activities from various papers into a holistic and structured collection. Managers of PSs at manufacturing companies can use this overview to reflect upon their existing approaches and identify overlooked activities worth to consider. These research findings offer several future research opportunities. First, the activities should be tested quantitatively. That is, linking the activities to the success of PSs to test if certain activities are more decisive for successfully managing PSs than others. Second, different combinations of activities might yield a positive outcome, as there might be no one-best-way to manage PSs. Thus, various combinations of activities should be tested instead of just a single one. Third, not all activities might be important when a PS is being designed, deployed or sustained. This implies that activities might increase or decrease in importance depending on the stage of PS implementation. Future research should pay more attention to these dynamics in a PS. Lastly, the collection of activities should be challenged and more refined by conducting interviews with more companies also from other industries.

References

[1] Womack, J.P., Jones, D.T., Roos, D., 1990. The machine that changed the world: How Japan's secret weapon in the global auto wars will revolutionize Western industry. HarperPerennial, New York, NY.

[2] Dombrowski, U., Mielke, T., 2015. Ganzheitliche Produktionssysteme. Springer Berlin Heidelberg, Berlin, Heidelberg.

[3] Netland, T., Ferdows, K., 2014. What to Expect From a Corporate Lean Program. MIT Sloan Management Review 55 (4), 83–89.

[4] Netland, T.H., 2014. Coordinating Production Improvement in International Production Networks: What's New?, in: Johansen, J., Farooq, S., Cheng, Y. (Eds.), International Operations Networks, vol. 29. Springer London, London, pp. 119–132.

[5] Bessant, J., Burnell, J., Harding, R., Webb, S., 1993. Continuous improvement in British manufacturing. Technovation 13 (4), 241–254.

[6] Bessant, J., Francis, D., 1999. Developing strategic continuous improvement capability. International Journal of Operations & Production Management 19 (11), 1106–1119.

[7] Poksinska, B., Swartling, D., 2018. From successful to sustainable Lean production – the case of a Lean Prize Award Winner. Total Quality Management & Business Excellence 29 (9-10), 996–1011.

[8] Schonberger, R.J., 2007. Japanese production management: An evolution-With mixed success. Journal of Operations Management 25 (2), 403–419.

[9] Sturdevant, D., 2014. (Still) learning from Toyota. McKinsey. https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/still-learning-from-toyota.

[10] Netland, T.H., 2013. Exploring the phenomenon of company-specific production systems: one-best-way or own-best-way? International Journal of Production Research 51 (4), 1084–1097.

[11] Netland, T.H., Aspelund, A., 2013. Company-specific production systems and competitive advantage. International Journal of Operations & Production Management 33 (11/12), 1511–1531.

[12] Dombrowski, U., Palluck, M., Schmidt, S., 2006. Typologisierung Ganzheitlicher Produktionssysteme. Zeitschrift für wirtschaftlichen Fabrikbetrieb 101 (10), 553–556.

[13] Jones, D.T., Womack, J.P., 2017. The Evolution of Lean Thinking and Practice, in: Powell, D.J., Netland, T.H. (Eds.), The Routledge companion to lean management. Routledge, New York, London, pp. 3–8.

[14] Dombrowski, U., Richter, T., Krenkel, P., 2017. Interdependencies of Industrie 4.0 & Lean Production Systems: A Use Cases Analysis. Procedia Manufacturing 11, 1061–1068.

[15] VDI - Verein Deutscher Ingenieure, 2012. Ganzheitliche Produktionssysteme.

[16] Netland, T.H., 2012. Managing strategic improvement programs: the XPS program management framework. Journal of Project, Program & Portfolio Management 3 (1), 31–44.

[17] Pellegrinelli, S., 1997. Programme management: organising project-based change. International Journal of Project Management 15 (3), 141–149.

[18] Hekneby, T., Ingvaldsen, J.A., Benders, J., 2020. Managing adoption by cultural development: Exploring the plant level effect of a 'Company Specific Production System' (XPS) in a Norwegian multinational company. Journal of Industrial Engineering and Management 13 (2), 402–416.

[19] Stalberg, L., Fundin, A., 2016. Exploring a holistic perspective on production system improvement. International Journal of Quality & Reliability Management 33 (2), 267–283.

[20] Netland, T.H., Aspelund, A., 2014. Multi-plant improvement programmes: a literature review and research agenda. International Journal of Operations & Production Management 34 (3), 390–418.

[21] Powell, D., Coughlan, P., 2020. Corporate Lean Programs: Practical Insights and Implications for Learning and Continuous Improvement. Procedia CIRP 93 (4), 820–825.

[22] Netland, T.H., Ferdows, K., Sanchez, E., 2015. How Company-Specific Production Systems Affect Plant Performance: The S-Curve Theory. Production and Operations Management 24 (3), 362–364.

[23] Stalberg, L., Fundin, A., 2018. Lean production integration adaptable to dynamic conditions. Journal of Manufacturing Technology Management 29 (8), 1358–1375.

[24] Saunders, M., Mann, R., Smith, R., 2008. Implementing strategic initiatives: a framework of leading practices. International Journal of Operations & Production Management 28 (11), 1095–1123.

[25] Thomé, A.M.T., Scavarda, L.F., Scavarda, A.J., 2016. Conducting systematic literature review in operations management. Production Planning & Control 27 (5), 408–420.

[26] Webster, J., Watson, R.T., 2002. Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly 26 (2), xiii–xxiii.

[27] Maritan, C.A., Brush, T.H., 2003. Heterogeneity and transferring practices: implementing flow manufacturing in multiple plants. Strategic Management Journal 24 (10), 945–959.

[28] Friedli, T., Bellm, D., 2013. Barriers and Success Factors in Managing Operational Excellence, in: Friedli, T., Basu, P., Bellm, D., Werani, J. (Eds.), Leading Pharmaceutical Operational Excellence. Outstanding Practices and Cases. Springer, Berlin, Heidelberg, pp. 103–114.

[29] Friedli, T., Werani, J., 2013. The History of OPEX in the Pharmaceutical Industry, in: Friedli, T., Basu, P., Bellm, D., Werani, J. (Eds.), Leading Pharmaceutical Operational Excellence. Outstanding Practices and Cases. Springer, Berlin, Heidelberg, pp. 27–34.

[30] Kvale, S., 1996. InterViews: An Introduction to Qualitative Research Interviewing. SAGE Publications, Thousand Oaks.

[31] Saldaña, J., 2013. The coding manual for qualitative researchers, 2. ed. ed. SAGE Publications, Los Angeles, CA.

[32] Glaser, B.G., Strauss, A.L., 1967. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Publishing, Chicago.

[33] Brown, A., 1992. Industrial experience with total quality management. Total Quality Management 3 (2), 147–156.

[34] Lee, B.-H., Jo, H.-J., 2007. The mutation of the Toyota Production System: adapting the TPS at Hyundai Motor Company. International Journal of Production Research 45 (16), 3665–3679.

[35] Bateman, N., David, A., 2002. Process improvement programmes: a model for assessing sustainability. International Journal of Operations & Production Management 22 (5), 515–526.

[36] Demeter, K., Losonci, D., 2019. Transferring lean knowledge within multinational networks. Production Planning & Control 30 (2-3), 211–224.

[37] Knol, W.H., Slomp, J., Schouteten, R.L.J., Lauche, K., 2018. Implementing lean practices in manufacturing SMEs: testing 'critical success factors' using Necessary Condition Analysis. International Journal of Production Research 56 (11), 3955–3973.

[38] Cagliano, R., Spina, G., 2000. How improvement programmes of manufacturing are selected. International Journal of Operations & Production Management 20 (7), 772–792.

[39] Paipa-Galeano, L., Bernal-Torres, C.A., Agudelo Otálora, L.M., Jarrah Nezhad, Y., González-Blanco, H.A., 2020. Key lessons to maintain continuous improvement: A case study of four companies. Journal of Industrial Engineering and Management 13 (1), 195–211.

[40] Sunder, V., Prashar, A., 2020. Empirical examination of critical failure factors of continuous improvement deployments: stage-wise results and a contingency theory perspective. International Journal of Production Research, 1–22.

[41] Netland, T.H., Ferdows, K., 2016. The S-Curve Effect of Lean Implementation. Production and Operations Management 25 (6), 1106–1120.

[42] Tortorella, G.L., Narayanamurthy, G., Thurer, M., 2021. Identifying pathways to a high-performing lean automation implementation: An empirical study in the manufacturing industry. International Journal of Production Economics 231 (3), 107918.

[43] Boscari, S., Danese, P., Romano, P., 2016. Implementation of lean production in multinational corporations: A case study of the transfer process from headquarters to subsidiaries. International Journal of Production Economics 176, 53–68.

[44] Danese, P., Romano, P., Boscari, S., 2017. The transfer process of lean practices in multi-plant companies. International Journal of Operations & Production Management 37 (4), 468–488.

[45] Keating, E., Oliva, R., Repenning, N., Rockart, S., Sterman, J., 1999. Overcoming the Improvement Paradox. European Management Journal 17 (2), 120–134.

[46] Prajogo, D., Sohal, A., 2004. The Sustainability and Evolution of Quality Improvement Programmes – an Australian Case Study. Total Quality Management & Business Excellence 15 (2), 205–220.

[47] Knol, W.H., Slomp, J., Schouteten, R.L.J., Lauche, K., 2019. The relative importance of improvement routines for implementing lean practices. International Journal of Operations & Production Management 39 (2), 214–237.

[48] Bhamu, J., Sangwan, K.S., 2016. A framework for lean manufacturing implementation. International Journal of Services and Operations Management 25 (3), 313–333.

[49] Secchi, R., Camuffo, A., 2016. Rolling out lean production systems: a knowledge-based perspective. International Journal of Operations & Production Management 36 (1), 61–85.

[50] Anand, G., Kodali, R., 2010. Development of a framework for implementation of lean manufacturing systems. International Journal of Management Practice 4 (1), 95–116.

[51] Done, A., Voss, C., Rytter, N.G., 2011. Best practice interventions: Short-term impact and long-term outcomes. Journal of Operations Management 29 (5), 500–513.

[52] Netland, T.H., 2016. Critical success factors for implementing lean production: the effect of contingencies. International Journal of Production Research 54 (8), 2433–2448.

[53] Hines, P., Taylor, D., Walsh, A., 2018. The Lean journey: have we got it wrong? Total Quality Management & Business Excellence 31 (3-4), 389–406.

[54] Inkpen, A.C., 2008. Knowledge transfer and international joint ventures: the case of NUMMI and General Motors. Strategic Management Journal 29 (4), 447–453.

[55] Turesky, E.F., Connell, P., 2010. Off the rails: understanding the derailment of a lean manufacturing initiative. Organization Management Journal 7 (2), 110–132.

[56] Galeazzo, A., Furlan, A., Vinelli, A., 2017. The organizational infrastructure of continuous improvement – an empirical analysis. Operations Management Research 10 (1-2), 33–46.

[57] Glover, W.J., Farris, J.A., van Aken, E.M., Doolen, T.L., 2011. Critical success factors for the sustainability of Kaizen event human resource outcomes: An empirical study. International Journal of Production Economics 132 (2), 197–213.

[58] Netland, T.H., Schloetzer, J.D., Ferdows, K., 2015. Implementing corporate lean programs: The effect of management control practices. Journal of Operations Management 36 (1), 90–102.

[59] Kim, Y.H., Sting, F.J., Loch, C.H., 2014. Top-down, bottom-up, or both? Toward an integrative perspective on operations strategy formation. Journal of Operations Management 32 (7-8), 462–474.

[60] Costa, F., Lispi, L., Staudacher, A.P., Rossini, M., Kundu, K., Cifone, F.D., 2019. How to foster Sustainable Continuous Improvement: A cause-effect relations map of Lean soft practices. Operations Research Perspectives 6 (1), 100091.

[61] Leonard, D., Reid, R., McAdam, R., 2002. A grounded multi-model framework for TQM dynamics. International Journal of Quality & Reliability Management 19 (6), 710–736.

[62] Colazo, J., 2021. Changes in communication patterns when implementing lean. International Journal of Quality & Reliability Management 38 (1), 296–316.

[63] Galeazzo, A., Furlan, A., Vinelli, A., 2021. The role of employees' participation and managers' authority on continuous improvement and performance. International Journal of Operations & Production Management 41 (13), 34–64.

Biography

Mark Grothkopp (*1994) has been a research associate in the operational excellence group at the Institute of Technology Management (ITEM) in the division of production management of the University of St.Gallen since 2019. He previously studied Industrial Engineering (M.Sc) at the Leuphana University Lüneburg.

Matteo Bernasconi (*1994) has been a research associate in the operational excellence group at the ITEM in the division of production management of the University of St.Gallen since 2020. He previously studied business innovation (M.A., B.A.) at the University of St.Gallen.

Prof. Dr. Thomas Friedli (*1971) is a director at the ITEM and became professor in 2004. He leads a division of 13 PhD students and two post-docs. His research areas include strategic management of production companies, management of industrial services, and operational excellence.

3rd Conference on Production Systems and Logistics

The Potential Of AutoML For Demand Forecasting

Kathrin Julia Kramer[1], Niclas Behn[2], Matthias Schmidt[1]
[1]*Institute of Product and Process Innovation (PPI)/ Leuphana University, Lueneburg, Germany*
[2]*Schaeffler Technologies AG & Co. KG, Schweinfurt, Germany*

Abstract

In demand forecasting, which can depend on various internal and external factors, machine learning (ML) methods can capture complex patterns and enable precise forecasts. Accurate forecasts facilitate targeted, demand-oriented planning and control of production and underline the importance of this task. The implementation of ML-algorithms requires knowledge of the specific domain as well as knowledge of data science and involves an elaborate set up process. This often makes the application of ML to potential industrial problems economically unattractive. The major skills shortage in the field of data science further exacerbates this. Automation and better accessibility of ML methods is therefore a key prerequisite for widespread use. This is where the principle of automated ML (AutoML) comes in, automating large parts of a ML pipeline and thus leading to a reduction in human labour input. Therefore, the aim of the publication is to investigate the extent to which AutoML solutions can generate added value for demand planning in the context of production planning and control. For this purpose, publicly available datasets deriving from Walmart as well as an anonymised manufacturing company are used for short-term and long-term forecasting. The AutoML tools from Microsoft, Dataiku and Google conduct these forecasts. Statistical models serve as benchmarks. The results show that the forecasting quality varies depending on the software, the input data and their demand patterns. Overall, the prepared models from Microsoft show the most accurate results in average and the potential of AutoML becomes particularly clear in the short-term forecast. This paper enriches the research field through its broad application, giving valuable insights into the use of AutoML tools for demand planning. The resulting understanding of limitations and benefits of AutoML tools for the case studies presented fosters their suitable application in practice.

Keywords

AutoML; Demand Forecasting; Sales Forecast; Machine Learning; Manufacturing; Production Planning

1. Introduction

ML-based demand forecasting offers the possibility to reflect various influencing parameters (e.g. currency exchange rates, sales region) and use those to identify complex non-linear patterns for forecasting future demand [1]. Accurate predictions enable an adequate planning of the production and procurement processes leading to less waste of resources (material, labour, capital) [2]. To foster the widespread use of machine learning (ML) in an industrial context, automated machine learning (AutoML) gains in importance as it aims to reduce the required knowledge in data science and time spend to set up a ML model [3–5]. Thus, it presents a possible solution for overcoming the skill shortage in the field of data science, which is currently one of the major barriers for the application of ML [2,6,7]. In addition, the shorter development time achieved by automating parts of a ML pipeline makes ML solutions more economically attractive [4,8]. Thus, researchers

DOI: https://doi.org/10.15488/12162

investigated the benefit of AutoML solutions in different case studies across various use cases (e.g. other production planning and control (PPC) tasks [9] or healthcare [10]). However, studies in regards to demand planning [11–13] in the context of PPC [14] remains a research gap. This paper aims at closing this gap by answering the research question: "*Can AutoML solutions support demand forecasting?*". The research question is embedded in the current state of research by highlighting the results of topic-related case studies. The third section presents the research methodology for achieving the objective of this paper. A generally valid answer to the research question is not the aim of this work, although it should be possible to make an assessment as broad as possible. To facilitate a comprehensive assessment, three AutoML tools in two production environments for forecasting demand of different product groups as well as for a short-term and long-term horizon are tested. Broad application enables a comprehensive evaluation of the prediction accuracy of the chosen AutoML tools and allows first conclusions for the use of AutoML solutions in demand forecasting. The fourth section presents the results of the research methodology used. The final section of the paper draws a conclusion and presents a future research agenda.

In summary, this paper enhances the current research through its broad application and comprehensive evaluation and thus allows conclusions on the potential use of AutoML tools for demand planning. In particular, an automated approach is required for scaling ML [8], i.e. ML-based prediction across different levels of observation (e.g. total, product group, product) and on a horizontal axis (e.g. each product).

2. Current state of research

This section explains the concept of AutoML and presents existing case studies that deal with AutoML in demand planning.

The herein used definition of AutoML was first introduced in 2014 and foresees an automation across the ML-pipeline [15]. The concept of AutoML has the objective to reduce the outlay of data scientists for ML projects and should instead enable domain experts to use ML methods without high level of statistical and ML knowledge [16]. More precisely, Yao et al. define AutoML as a combination of automation and ML and understand AutoML as an automated setup of a ML-pipeline with limited computing power and limited (or no) human support [17]. A common ML-pipeline consists of business and data understanding, data preparation, modelling, evaluation and deployment tasks, e.g. used in the industry-independent Cross Industry Standard Process for Data Mining (CRISP-DM) [18]. This framework is already used for demand forecasting [19–21]. The nature of the process involves feedback loops and continuous readjustments of assumptions, forming a life cycle process [18]. In particular, the processes of data understanding, data preparation, modelling and evaluation require the expertise of data scientists and could therefore be automated [5]: data preparation foresees to select features, clean and transform those as well as generate new features. The modelling process contains the choice of an algorithm and the optimization of its hyper-parameters and for artificial neural networks (ANN) also the definition of the net architecture [18,5]. Various facets can be investigated during models' evaluation, whereby the prediction accuracy is usually the focus [22]. The evaluation can take place when the algorithm converges during training or in order to save time as well as computing power, for instance a predefined budget of computing resources can be used as stop criteria [5].

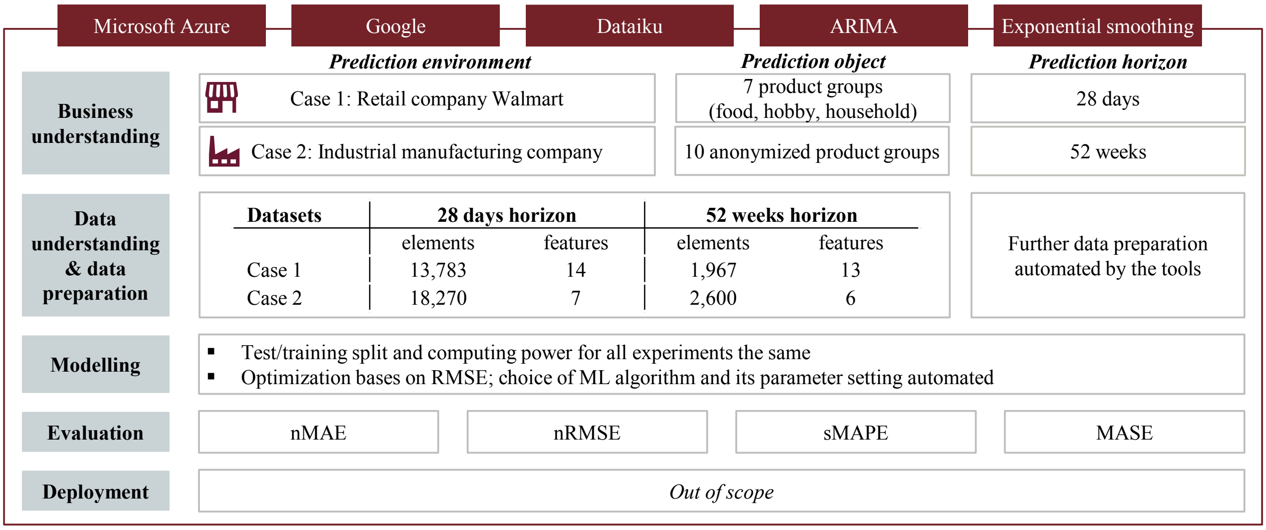
For demand planning, AutoML has already been used in several studies. The focus of these studies is primarily in the sales and marketing environment [11–13]. The study by Gonçalves et al. is the only contribution that investigates AutoML for demand planning in the context of PPC [14]. Ford et al. apply a self-developed AutoML solution for the products of an alcoholic beverage distributor. For this purpose, they generate forecasts for three horizons, 1 month, 3 months and 12 months, and include autoregressive methods and the average for a comparison. They use two univariate datasets and mean squared error as a quality criterion. The AutoML models achieve worse results, which are below the forecasting quality of the autoregressive methods. For one of the datasets, which is characterised by a high proportion of noise, the average method achieves the best results [12]. Henzel and Sikora use AutoML to create ANN and compare them with manually created XGBoost and ANN models. They use a dataset of everyday consumer goods and focus on the impact of promotions on sales. The models created with AutoML are superior to both manually created models in 2 of 12 cases (according to root mean squared error (RMSE) and mean absolute error (MAE)), and better than the manually created ANN in all cases [13]. Dai and Huang create a Long Short Term Memory (LSTM) model with a special loss function, which they optimise with hyper-parameter search. For comparison, they create six ML models with AutoML, which they use as benchmarks. They use a sparse consumer goods dataset that contains causal information of sales (e.g. holidays, promotions). The authors use weekly and monthly data to forecast cumulative sales at the store level. The LSTM model achieves better results (according to mean absolute percentage error (MAPE) and root mean squared percentage error) compared to the AutoML models, whereby the performance of the LSTM model decreases with increasing forecast horizon [11]. Gonçalves et al. compare statistical (Naïve, exponential smoothing, ARIMA, ARIMAX) and ML models (feed-forward ANN, random forest, support vector regression, recurrent ANN), including a model with AutoML. They forecast the demand for electronic components of a manufacturer from the automotive supply sector and follow a multivariate approach with various leading indicators. In particular, they investigate whether and to what extent a multivariate approach is superior to a univariate approach in different phases of the product life cycle. In the quality criterion normalized MAE (nMAE) used, AutoML achieves the fourth best performance (out of 9) on average across all phases of the product life cycle, with a considerable gap between it and the following statistical methods: Naïve, exponential smoothing and ARIMA [14].

The studies presented show that AutoML models tend to perform worse than manually created ML models and better than statistical models. This is particularly the case with multivariate forecasts. For univariate forecasts and one-step forecasts, statistical methods tend to perform better. With the few identified studies, the described area of research is still underrepresented. Thus, further investigations in the field of demand forecasting, and especially in the context of PPC are necessary. To add to this note, continuous progress in the field of AutoML makes results of past studies hard to interpret for assessing the potential of current AutoML solutions. In addition, existing studies on demand forecasting have not yet compared several AutoML solutions. Thus, this paper contributes to the existing research by conducting a comparison of several AutoML solutions with statistical methods for the area of demand forecasting in the context of PPC.

3. Research methodology

The research methodology of this paper is of empirical nature. A transparent setup of different experiments enables an in-depth understanding of the limitations and benefits when using the chosen AutoML tools and facilitates researchers to transfer this methodology to different case studies as well as AutoML tools.

To begin, the authors identified 31 existing AutoML tools. The pool of potential tools reduces to eight tools as only those tools fulfil the following criteria: they offer a test version or academic licence, can handle time series data and offer at least partial automatic data preparation. Of these, three tools are chosen as examples for the investigations in this paper: the AutoML solutions Microsoft Azure Automated ML, Google Cloud AutoML Tables and Dataiku Data Science Studio. To emphasise again, a selection of more than one tool is of importance to understand possible deviations across the tools. An aspect that was so far not analysed. ARIMA and exponential smoothing are taken as benchmarks as these methods are most commonly used in practice and do not require extensive data science knowledge [23,24]. For the investigation of the chosen tools, the following assumptions are considered with the objective to facilitate a comparable set up as well as test the AutoML solutions in different settings. Figure 1 shows these specifications that are structured according to the CRISP-DM phases (grey boxes).



Datasets	28 days horizon		52 weeks horizon	
	elements	features	elements	features
Case 1	13,783	14	1,967	13
Case 2	18,270	7	2,600	6

- Test/training split and computing power for all experiments the same
- Optimization bases on RMSE; choice of ML algorithm and its parameter setting automated

Figure 1: Specifications of examinations

The demand forecasts take place for two different companies: the retail goods company Walmart (case 1) and an anonymized company that manufactures industrial goods (case 2). In order to compare the forecast quality at different forecast horizons, a yearly and monthly forecast will be prepared [25]. 28 days and 52 weeks are chosen as prediction horizons to analyse whether there are performance differences between a short- and long-term perspective. The aim is to obtain the total demand per product group (label attribute) in order to support decision-making processes, e.g. the decision on long-term supplier contracts and the acquisition of a new production plant from a long-term perspective or the timing of the procurement of bought-in parts and production from a medium-term perspective. To investigate how robust the AutoML solutions predict, it is of importance to test the tools on different demand patterns. This is the case for the chosen product groups, seven product groups for case 1 from the food, hobby and household sectors and ten anonymized product groups for case 2. The datasets are publicly available [26,27]. As the PPC is the focus of this work, the public datasets have been modified to remove the store/ warehouse levels. The datasets consist of 13,783 elements (case 1) and 18,270 elements (case 2) of daily demand per product group for the prediction horizon of 28 days. For the prediction horizon of 52 weeks, case 1 has 1,967 elements and case 2 counts 2,600 elements of weekly demand per product group. Table 1 gives an overview of the different groups of features. Three main differences exist regarding the features. First, the dataset of case 1 counts more features that are descriptive then case 2. It contains additional information on the occurrence of public events and governmental support schemes of three US states. The dataset of case 2 lacks these features and only consists of the label attribute and the date. To test the tools on a multivariate setting, further time information (e.g. weekday, calendar week, year) are added. Secondly, when transforming the datasets from daily to weekly data, some features are excluded (e.g. weekday), transformed (e.g. event features) or added (e.g. calendar week). Thirdly, depending on the AutoML tool, further features are added: the ML tools of Microsoft Azure and Dataiku add information on school and bank holidays. This addition is only made in case 1, as the region is known, whereas it is unknown in case 2. Besides, Microsoft Azure Automated ML is the only tool that generates features about the seasonal and trend component of the demand time series as well as lag features for public events. Thus, this tool generates most features in comparison to the other two tools.

Table 1: Overview of features

Initial feature group	**Case 1**			**Case 2**		
Demand per product group (label)	x			x		
Time information (e.g. date, week)	x			x		
Public events	x					
Governmental support scheme	x					
Additional AutoML feature group	**Azure**	**Dataiku**	**Google**	**Azure**	**Dataiku**	**Google**
Bank- and school holidays	x	x				
Lag features for public events	x					
Seasonal and trend components	x			x		
Number of reflected feature groups	**7**	**5**	**4**	**3**	**2**	**2**

x = part of the dataset Azure = AutoML tool by Microsoft Azure Dataiku = AutoML tool by Dataiku Google = AutoML tool by Google

The dataset of case 2 contained missing values that were replaced by zero to enable a proper use of the tools. Besides that, the selected AutoML tools perform the remaining data preparation (e.g. feature selection, feature generation, data normalization). By means of this paper, the objective is to investigate the performance of the automated ML processes. Thus, no further manual preparation takes place. Modelling bases also on the automated decisions of the tools. However, for facilitating comparable results, the test split corresponds to the forecasting horizon, optimization bases on RMSE and on a constant computing power available for training across all tools. For evaluation of the models, first it is outlined whether AutoML tools are able to create predictions that are better than Naïve forecasts and secondly, if those models perform better than the benchmark of statistical methods. The evaluation metrics nMAE, nRMSE, sMAPE, and mean absolute scaled error (MASE) are selected. MAE is to be used because it is an absolute and scale-based measure that has been used before in similar research projects. RMSE is also frequently used [13,24,28]. This paper uses the normalised version of MAE and RMSE (nMAE and nRMSE) so that a comparison over several time series is possible. The mean of the actual values of the forecast horizon is chosen for normalisation. The symmetrical variant of the mean absolute percentage error sMAPE is used as a percentage quality measure. This is more robust than MAPE and allows the evaluation of zero values. Both quality measures are frequently used in the evaluation of time series [24,29]. Hyndman & Koehler argue that the value of sMAPE can be unstable and instead recommend the use of MASE [30]. Thus, the last measure used is the relative quality measure MASE, which has been applied in several studies [23,24,30,31]. The Naïve method functions as a benchmark model for MASE. If the calculated value is below 1 the forecast is better than a Naïve forecast and vice versa [30]. The final phase, the deployment of models, exceeds the scope of this paper.

Overall, the research methodology enables a comprehensive analysis of the potential from AutoML tools with regard to the prediction accuracy for the chosen use cases. By looking at different datasets with various demand patterns and different input features as well as multiple evaluation metrics, this paper contributes to the research field. The results help to understand the limitations and potentials of the observed tools in the investigated environments and lead to hypotheses for future applications.

4. Results

This section presents the results of the applied research methodology. The results relate to the best performing and thus chosen model of each AutoML tool as well as ARIMA and exponential smoothing. The training of the respective models took 42 minutes to 100.2 minutes. The first part of the analysis is to check whether the prediction accuracy of the best performing model according to MASE per product group of case 1 and case 2 is above or below 1 for both prediction horizons. In case 1, the predictions of at least one model across all product groups are better than a Naïve forecast. However, in case 2, the provided predictions for product group 001, 007 and 011 for the 52-weeks horizon do not show sufficient results as they are as good as a Naïve forecast. As second part of the analysis, Table 2 summarizes the average results of the best AutoML tool in comparison to the best statistical model. For the following comparison, the product groups 001, 007 and 011 are disregarded as they would otherwise distort the picture.

Table 2: Prediction accuracy of the best AutoML tool in comparison to best statistical model

Case	**Horizon**	**Best AutoML tool**	**Best statistical method**	**nMAE**	**nRMSE**	**sMAPE**	**MASE**
1	28 days	Azure	Exponential smoothing	+17.87%	+18.96%	+17.62%	+23.88%
	52 weeks	Dataiku	ARIMA	+2.92%	+7.99%	+2.11%	-4.46%
2	28 days	Azure	ARIMA	+18.81%	+17.81%	+8.05%	+7.11%
	52 weeks	Dataiku* / Azure**	Exponential smoothing	-10.71%	-10.08%	-6.51%	-10.85%

* according to nMAE & nRMSE ** according to sMAPE & MASE

It shows that the prediction accuracy of AutoML across the different experiments is in average in particular beneficial for the short-term forecast. The best performing AutoML tool in this case is Microsoft Azure's tool. Table 2 displays, depending on the evaluation metric, an average improvement of 17.62% to 23.88% for the first case study and 7.11% to 18.81% for the second case study in comparison to the best performing statistical model. The long-term forecasting reflects mixed results. For case 1, three of four evaluation metrics reflect an improved prediction of AutoML by 2.11% to 7.99% in comparison to a statistical method. However, the MASE value indicates a negative effect of -4.46%. The results of the best performing AutoML model in case 2 show worse results (-6.51% to -10.85%) in contrast to the best performing statistical model. To get a more detailed picture of the prediction accuracy of each AutoML tool as well as the benchmark of statistical models across the different product groups, the following figure illustrates the distribution through boxplots of each model according to MASE.

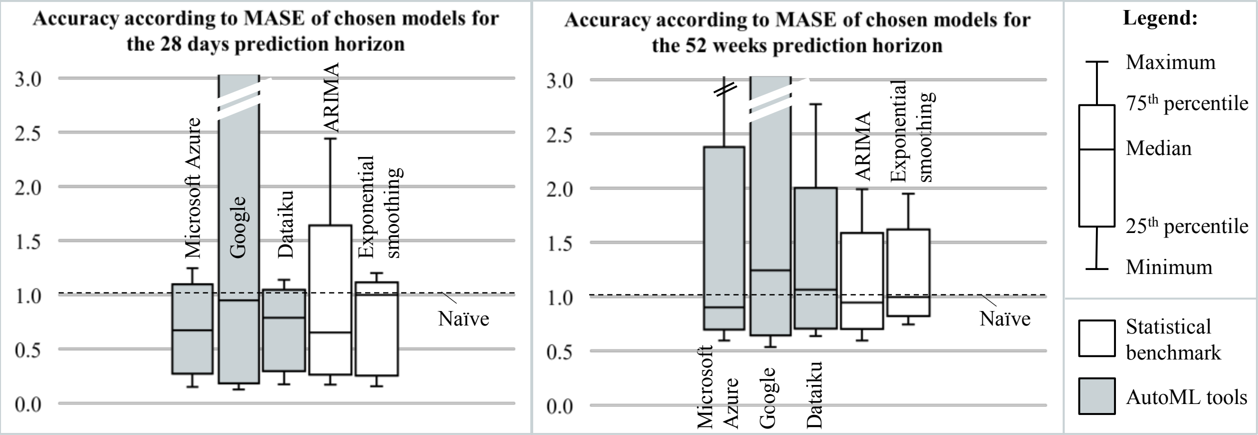


Figure 2: Box plot diagram of prediction accuracies from all product groups per model

It shows that the results of the AutoML tool by Google varies the most. In one case the tool receives the best prediction accuracy (case 1, product group Food 3, prediction horizon of 28 days with MASE of 0.127) and in another case the worst accuracy (case 2, product group Category_001, prediction horizon of 28 days with MASE of 33.885) in relation to all product groups. For the other tools, differences between the two prediction horizons exist. On the short-term horizon, the tool of Dataiku varies the least (MASE between 0.175-1.139), closely followed by exponential smoothing (MASE between 0.156-1.200) and Microsoft Azure's tool (MASE between 0.152-1.246). ARIMA has the lowest median with a MASE value of 0.651, but the second most variation of all values. The tool of Microsoft Azure points out the second lowest median. On the long-term prediction horizon, Figure 2 shows that the prediction accuracy's variation of the statistical models ARIMA and exponential smoothing is lower than the variation from the AutoML models. However, only the median of ARIMA (median MASE of 0.945) and Microsoft Azure (median MASE of 0.901) was lower than one and is thus better than a Naïve forecast. Please refer to Table 3 for an overview of key statistics, also of the other evaluation metrics. When ranking the model according to prediction accuracy, in some cases the ranking differs when looking at different evaluation metrics. For example, the best performing model for the long-term horizon, thus the minimum value of considered metric, is in the case of MAE, nRMSE, sMAPE the ARIMA model and for MASE the tool by Google. However, when looking at the associated mean across all product groups of both case studies, ARIMA is the model in favour.

Table 3: key statistics of investigated models across all product groups of both case studies

	Expon. smoothing	**ARIMA**	**Microsoft Azure**	**Dataiku**	**Google**	**Expon. smoothing**	**ARIMA**	**Microsoft Azure**	**Dataiku**	**Google**	**Expon. smoothing**	**ARIMA**	**Microsoft Azure**	**Dataiku**	**Google**	**Expon. smoothing**	**ARIMA**	**Microsoft Azure**	**Dataiku**	**Google**
28 days	MAE					nRMSE					sMAPE					MASE				
Min.	0.046	0.043	0.041	0.043	0.031	0.054	0.052	0.050	0.054	0.040	0.047	0.043	0.042	0.044	0.031	0.156	0.173	0.152	0.175	0.127
1st Q.	0.058	0.059	0.065	0.063	0.071	0.081	0.081	0.079	0.085	0.085	0.056	0.057	0.069	0.062	0.075	0.352	0.355	0.390	0.418	0.240
Median	0.193	0.151	0.189	0.171	0.205	0.241	0.204	0.232	0.195	0.250	0.196	0.147	0.189	0.168	0.214	1.000	0.651	0.672	0.790	0.948
3rd Q.	0.311	0.324	0.294	0.295	2.969	0.402	0.408	0.344	0.373	3.646	0.324	0.353	0.315	0.313	1.874	1.029	0.837	0.952	0.952	5.463
Max.	1.092	0.756	0.544	0.812	22.81	1.181	0.907	0.591	0.887	29.97	0.760	0.582	0.498	0.641	2.000	1.200	2.441	1.246	1.139	33.88
Mean	0.279	0.236	0.202	0.230	2.825	0.338	0.283	0.244	0.275	3.664	0.248	0.223	0.206	0.226	0.681	0.768	0.722	0.662	0.682	5.932
St. Dev.	0.294	0.218	0.147	0.204	5.983	0.325	0.249	0.171	0.228	7.868	0.219	0.184	0.149	0.186	0.845	0.391	0.560	0.394	0.338	9.933
52 weeks	MAE					nRMSE					sMAPE					MASE				
Min.	0.047	0.032	0.065	0.061	0.059	0.055	0.041	0.079	0.074	0.073	0.047	0.033	0.065	0.063	0.059	0.743	0.595	0.596	0.638	0.537
1st Q.	0.099	0.094	0.093	0.074	0.107	0.123	0.123	0.112	0.095	0.131	0.101	0.096	0.093	0.075	0.110	0.897	0.810	0.796	0.773	0.750
Median	0.213	0.169	0.163	0.186	0.174	0.241	0.242	0.200	0.204	0.200	0.209	0.170	0.157	0.173	0.183	0.996	0.945	0.901	1.064	1.242
3rd Q.	0.259	0.321	0.227	0.295	0.858	0.309	0.364	0.266	0.398	1.187	0.243	0.290	0.225	0.287	0.730	1.290	1.185	1.102	1.230	2.427
Max.	0.845	0.514	1.163	0.638	4.214	0.888	0.579	1.176	0.674	5.982	0.942	0.453	0.744	0.499	1.470	1.948	1.990	3.654	2.774	17.04
Mean	0.245	0.222	0.247	0.236	0.778	0.284	0.263	0.287	0.274	1.074	0.244	0.206	0.218	0.216	0.439	1.111	1.033	1.109	1.124	3.472
St. Dev.	0.214	0.153	0.275	0.177	1.300	0.234	0.171	0.283	0.191	1.838	0.226	0.131	0.187	0.144	0.493	0.345	0.367	0.699	0.527	5.300

Looking at the short-term prediction in most cases no big differences occur when comparing the average prediction accuracy of AutoML with the statistical benchmark. In 10 out of 17 product groups (58.8%) the average deviation equals to MASE of -0.02-0.18. The remaining seven product groups vary between -0.36 and 11.08. The deviation is even lower when comparing the best performing model. There, only the product group Food 2 and 028 show a deviation of -0.53 and -0.76. Illustrative, the left side of Figure 3 presents the predicted demand of the best performing AutoML model (line-dotted line) and statistical model (circular-dotted line) in relation to the actual demand of product group 028. The deviation of the remaining product groups is only -0.14-+0.14.

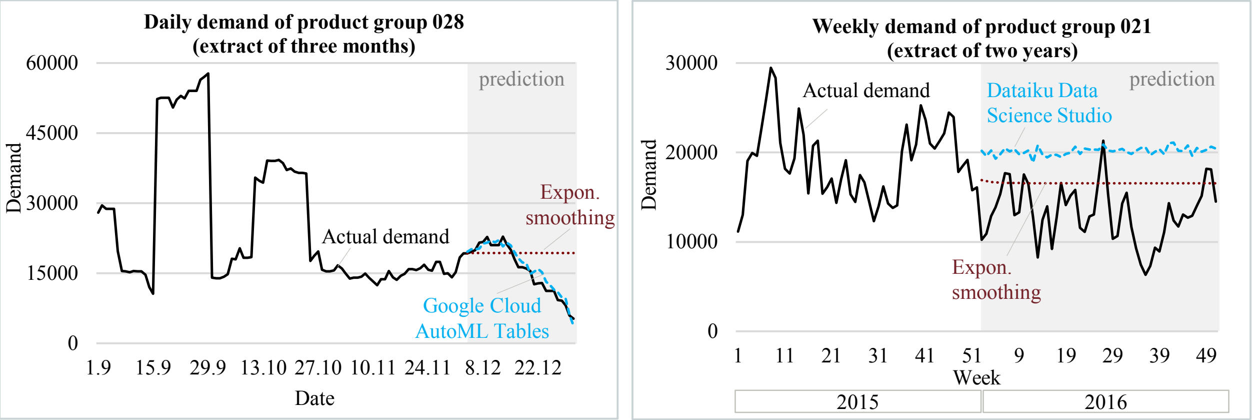


Figure 3: Extract of predicted demand from selected models and actual demand

For the long-term prediction the average deviation between AutoML and the statistical benchmark is in 10 out of 17 product groups between -0.31-+0.25. For the remaining seven product groups, five groups in favour of statistical models (in average 0.50-5.36 more precise than AutoML) and two groups in favour of AutoML (in average 0.68 and 0.85 more precise than the statistical benchmark). When looking at most precise model and not the average, the best performing model of six product groups deviates more than 0.5 between statistical and AutoML models: four in favour for AutoML (0.52-0.60 more precise than the best performing statistical model) and two in favour of the best performing statistical model (0.56 and 0.84 more precise than the best AutoML model). Product group 021, shown in Figure 3, corresponds to the biggest deviation in favour of the statistical benchmark, when looking at the MASE values. Thus, Figure 3 presents on the left side the biggest difference in favour of AutoML and the right side in favour of a statistical model.

5. Conclusion and future research agenda

The results show a mixed picture. On the one hand, some AutoML tools perform well in certain scenarios and far better than statistical models in a few cases. However, in other scenarios AutoML is less accurate than statistical models, even in certain scenarios by far. AutoML, namely the tools by Dataiku and Microsoft Azure, tend to predict more stable in short-term predictions, but ARIMA has a slightly lower median of MASE than Dataiku across all product groups. Especially for case 1, AutoML achieves for six of seven product groups the lowest MASE value. This could be due to the fact that descriptive features are part of the dataset. For case 2, where only additional time information exists, only in three out of ten product groups, an AutoML tool was prior to the rest. However, in average all models are performing worse than in case 1. The results of the long-term predictions for case 1 show that at least one AutoML tool performs best for four groups and that a statistical model is the preferred choice for the other three groups. The assumed advantage through more descriptive feature is not as clear as for the short-term prediction horizon. In case 2, at least one AutoML model predicts the demand of most product groups (6 out of 10) more precisely than a statistical model. However, the results of ARIMA and exponential smoothing deviate not as much as of the AutoML tools. The AutoML tool of Google seems to be most sensitive to the balance of data as some product groups were more frequently demanded than others. Especially in case 2, prediction accuracies for less demanded product groups (e.g. 001, 011, 015, 021, 024) are far above a MASE value of 1. Nevertheless, Google's tool also trains the most accurate model with a MASE of 0.127 (product group Food 3). As this paper shows, AutoML can support on preparation tasks, modelling and the associated optimization of hyper-parameters as well as the evaluation of models. However, as this analysis expresses prediction results are not fully reliable yet. Further investigations should take place to understand the differences in performance. As a ML pipeline includes several assumptions, some aspects for further investigations are outlined: Firstly, modelling with more features should be conducted, to test the hypothesis that AutoML is in favour when predicting on a multivariate basis. The herein analysed datasets have only few features, which presumably explains why the models partially reach their limits when it comes to long-term forecasting. In addition, the test and training split should be varied to get further insights into an eventual over- or underfitting of a ML model. Also, to understand the sensitivity of AutoML tools further demand patterns, longer history of data, different settings of data preparation (e.g. keeping missing values), single and global models, prediction horizons and different case studies should be investigated. Moreover, further statistical models (e.g. ARIMAX), manually trained ML-models and fuzzy models could function as additional benchmarks that should be evaluated on prediction accuracy, running and implementation time. In summary, the results can help to find a sweet spot for the use of AutoML. This howsoever highly depends on the manufacturing setup, i.e. the required prediction horizon results e.g. from procurement time for materials, storage capacity and production lead times. It should be noted that the analyses are repeated regularly to examine the progress of AutoML/ ML over time.

Thus, in conclusion, the research question whether AutoML can support demand planning in PPC can be answered as following: this paper shows especially for short-term predictions good results for AutoML. However, for some demand patterns and less demanded product groups, the accuracy was not sufficient. Thus, AutoML can function for prototyping and can be part in business processes. When implementing into business processes the chosen AutoML tool should be regularly tested against a benchmark of different ML and statistical models. AutoML can significantly reduce the time spend to analyse the provided data as well as train and optimize different algorithms and can therefore be a first step for companies to test ML in their business environment. Therefore, it can help to ensure that domain expertise is effectively reflected in data-driven models by enabling domain experts to use ML without having extensive data science knowledge. Nevertheless, ML and AutoML comes at the cost of models that are more complex and use more computing power in comparison to statistical models [32]. With the use of AutoML tools the understanding of the ‘engine’ behind the models, for instance how the feature engineering or optimizing of the parameters from the algorithm take place, becomes less transparent as they are for most tools not visible in the user interface. In the future, aspects such as user-friendliness, transparency and trustworthiness of workflows should be considered next to the tools’ prediction accuracy, running and implementation time as well as robustness.

Acknowledgements

Funded by the Lower Saxony Ministry of Science and Culture under grant number ZN3489 within the Lower Saxony “Vorab” of the Volkswagen Foundation and supported by the Center for Digital Innovations (ZDIN).

References

[1] Gentsch, P., 2019. Künstliche Intelligenz für Sales, Marketing und Service, 2nd ed. Springer Fachmedien Wiesbaden, Wiesbaden.

[2] Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P., Malhotra, S., 2018. Notes from the AI frontier: Insights from hundreds of use-cases, 36 pp. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning. Accessed 16 July 2021.

[3] BMWi, 2015. Erschließen der Potenziale der Anwendung von "Industrie 4.0" im Mittelstand. Bundesministerium für Wirtschaft und Energie. https://www.bmwi.de/Redaktion/DE/Publikationen/Studien/erschliessen-der-potenziale-der-anwendung-von-industrie-4-0-im-mittelstand.html. Accessed 25 August 2021.

[4] Elshawi, R., Sakr, S., 2020. Automated Machine Learning: Techniques and Frameworks, in: Kutsche, R.-D., Zimányi, E. (Eds.), Big Data Management and Analytics. 9th European Summer School, eBISS 2019, Berlin, Germany, June 30 – July 5, 2019, Revised Selected Papers, vol. 390, 1st ed. 2020 ed. Springer International Publishing, Cham, pp. 40–69.

[5] He, X., Zhao, K., Chu, X., 2020. AutoML: A survey of the state-of-the-art. Knowledge-Based Systems.

[6] Deloitte. State of AI in the Enterprise - 3rd Edition. https://www2.deloitte.com/de/de/pages/technology-media-and-telecommunications/articles/ki-studie-2020.html. Accessed 17 January 2022.

[7] PwC. Künstliche Intelligenz in Unternehmen. https://www.pwc.de/de/digitale-transformation/kuenstliche-intelligenz/kuenstliche-intelligenz-in-unternehmen.html. Accessed 17 January 2022.

[8] Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated Machine Learning. Springer Intl Publishing, Cham.

[9] Bender, J., Ovtcharova, J., 2021. Prototyping Machine-Learning-Supported Lead Time Prediction Using AutoML. Procedia Computer Science 180, 649–655.

[10] Waring, J., Lindvall, C., Umeton, R., 2020. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial intelligence in medicine (104), 1–12.

[11] Dai, Y., Huang, J., 2021. A Sales Prediction Method Based on LSTM with Hyper-Parameter Search. J. Phys.: Conf. Ser. (1756).

[12] Ford, J., Nava, C., Tan, J., Sadler, B., 2020. Automated Machine Learning Framework for Demand Forecasting in Wholesale Beverage Alcohol Distribution. SMU Data Science Review (3).

[13] Henzel, J., Sikora, M., 2020. Gradient Boosting and Deep Learning Models Approach to Forecasting Promotions Efficiency in FMCG Retail, in: Rutkowski, L., Scherer, R., Korytkowski, M., Pedrycz, W., Tadeusiewicz, R., Zurada, J.M. (Eds.), Artificial Intelligence and Soft Computing, vol. 12416. Springer International Publishing, Cham, pp. 336–345.

[14] Gonçalves, J.N., Cortez, P., Carvalho, M.S., Frazão, N.M., 2020. A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain. Decision Support Systems (142).

[15] Hutter, F., Caruana, R., Bardenet, R., Bilenko, M., Guyon, I., Kegl, B., Larochelle H., 2014. AutoML 2014 workshop. https://sites.google.com/site/automlwsicml14/. Accessed 29 November 2021.

[16] Zöller, M.-A., Huber, M.F., 2021. Benchmark and Survey of Automated Machine Learning Frameworks. Journal of Artificial Intelligence Research (70), 409–472.

[17] Yao, Q., Wang, M., Chen, Y., Dai, W., Li, Y.-F., Tu, W.-W., Yang, Q., Yu, Y., 2018. Taking Human out of Learning Applications: A Survey on Automated Machine Learning, 20 pp. Accessed 29 November 2021.

[18] Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer, 2000. CRISP-DM 1.0 Step-by-step data mining guide.

[19] Alegado, R.T., Tumibay, G.M., 2020. Statistical and Machine Learning Methods for Vaccine Demand Forecasting: A Comparative Analysis. Journal of Computer and Communications (Vol.8), 37–49.

[20] Maaß, D., Spruit, M., Waal, P. de, 2014. Improving short-term demand forecasting for short-lifecycle consumer products with data mining techniques. Decis. Anal. 1 (1).

[21] Schreiber, L., Moroff, N., 2020. Machine Learning versus Statistical Methods in Demand Planning for Energy-Efficient Supply Chains. ICoMS 2020: Proceedings of the 2020 3rd International Conference on Mathematics and Statistics, 17–23.

[22] Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining: Practical machine learning tools and techniques, Third edition ed. Morgan Kaufmann, Amsterdam.

[23] Cerqueira, V., Torgo, L., Soares, C., 2019. Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters, 9 pp. http://arxiv.org/pdf/1909.13316v1. Accessed 25 May 2021.

[24] Hyndman, R.J., Athanasopoulos, G., 2021. Forecasting: principles and practice. OTexts.com/fpp3. Accessed 2 December 2021.

[25] Wiendahl, H.-P., Wiendahl, H.-H., 2019. Betriebsorganisation für Ingenieure, 9., vollständig überarbeitete Auflage ed. Hanser, München.

[26] FelixZhao, 2017. Forecasts for Product Demand: Make Accurate Forecasts for Thousands of Different Products. Data. https://www.kaggle.com/felixzhao/productdemandforecasting. Accessed 2 December 2021.

[27] University of Nicosia, 2020. M5 Forecasting - Accuracy: Estimate the unit sales of Walmart retail goods. data. https://www.kaggle.com/c/m5-forecasting-accuracy/data. Accessed 2 December 2021.

[28] Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY.

[29] Javeri, I.Y., Toutiaee, M., Arpinar, I.B., Miller, T.W., Miller, J.A., 2021. Improving Neural Networks for Time Series Forecasting using Data Augmentation and AutoML. http://arxiv.org/pdf/2103.01992v3. Accessed 30 June 2021.

[30] Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. International Journal of Forecasting 22 (4), 679–688.

[31] Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and Machine Learning forecasting methods: Concerns and ways forward. PloS one 13 (3).

[32] Russell, S.J., Norvig, P., 2016. Artificial intelligence: A modern approach, Third edition, Global edition ed. Pearson, Boston, Columbus, Indianapolis.

Biography

Kathrin Julia Kramer (*1991) is an Academic Assistant and PhD-student in the field of production management at the Institute of Product and Process Innovation (PPI) at the Leuphana University Lueneburg since 2019. Her research focus is on the practical application of machine learning methods in production planning and control.

Niclas Behn (*1993) is a Graduate Trainee in Operations. In 2021, he completed his master degree in Management & Engineering at the Institute of Product and Process Innovation (PPI) at the Leuphana University Lueneburg.

Matthias Schmidt (*1978) studied industrial engineering at the Leibniz University Hannover and subsequently worked as a research associate at the Institute of Production Systems and Logistics (IFA). After completing his doctorate in engineering, he became head of Research and Industry of the IFA and received his habilitation. Since 2018, he holds the chair of production management at the Institute for Product and Process Innovation (PPI) at the Leuphana University of Lueneburg. In addition, he became the head of the PPI in 2019.

3rd Conference on Production Systems and Logistics

Business Processes Digitalization as a Resolution Direction for Digital Operations Challenges in Digital Supply Networks

Andreas M. Radke[1], Thorsten Wuest[2] and David Romero[3]
[1] *mSE North America, Inc., Chicago IL 60611, USA*
[2] *West Virginia University, Morgantown, WV 26506, USA*
[3] *Tecnológico de Monterrey, Mexico City 14380, Mexico*

Abstract

Supply chain maturity models urge increased collaboration among supply chain participants to achieve sustained competitive advantage through operational end-to-end visibility and transparency. Digital solutions provide the tools and technologies to enable Digital Supply Networks through inter-connectivity among supply network partners. A major challenge in improving digitalized operations is the divergence between 'official' business processes that are mapped out with accountabilities assigned and the 'de facto' business processes that are executed. During business processes optimization and IT system enhancements, this mismatch between documented versus tribal business processes results often in inefficient, ineffective, and if not addressed early, infeasible digital solutions. In this paper, the authors outline challenges and their root causes by discussing possible resolution directions in the dimensions of organizational change management (e.g., connected customer), IT systems gaps (e.g., composable applications), and common datasets for digital operations and business process mining applications (e.g., digital twins).

Keywords

Business-IT Alignment; Business Processes (Re-)engineering; Digitalization; Digital Operations; Digital Supply Networks; Digital Technologies.

1. Introduction

To improve the agility and resilience to disruptions of supply network operations enterprise decision-makers across all industries continue to invest heavily in Information Technology (IT) [1]. From a technological perspective, it can be expected major improvements due to decades of exponential advances in data processing power (Moore's Law) [2], volume of data transferability (Butter's Law) [3, 4], and storage density per dollar (Kreider's Law) [5]. However, for several years, before being unveiled by the recent pandemic's disruptions, supply chain innovation and performance have deteriorated for a majority of enterprises according to industry observers [6]. Assessing the effect of COVID-19 on supply chain performance, one observation is the failure of integrated supply chain IT solutions to provide "agility" [6].

Supported by anecdotal evidence of supply chain operations, the single most important supply chain management system is Microsoft Excel for its ubiquitous availability, versatility, and speed of adaptability. Both practitioners and researchers, seem to focus on a technology solution to a problem that is neither well documented nor researched in-depth. This paper aims to highlight the gaps in research and practice for Business Process Digitalization efforts in the context of Digital Supply Networks.

DOI: https://doi.org/10.15488/12117 publish-Ing.

To achieve this objective, two guiding hypotheses are formulated:

- The available academic and grey literature focuses on interpreting supply networks' digitalization efforts as a technological challenge and provides frameworks and methodologies accordingly.
- There remain significant gaps in academic research and industrial practice to address the interplay between business processes and (digital) technologies (i.e., business-IT alignment).

To address these hypotheses, this paper provides a background on the state-of-the-art of Digital Supply Networks in Section 2. The research and practice gaps, as well as current approaches to bridge business processes and digital technologies, are reviewed in Section 3. Section 4 presents practical use cases for common pitfalls in business processes and digital technologies integration efforts. Section 5 provides discussion and research agenda, and the paper concludes in Section 6 with final reflections.

2. Background

Digital Transformation is on the top of the priority list for most manufacturing companies and their supply networks. Many are overwhelmed and struggling with the transition and tend to approach it mainly from a technological perspective (aka. 'can we do it' versus 'should we do it'). Selecting an appropriate reference architecture for information systems integration and/or interoperability, and enabling frameworks and technologies (e.g., cybersecurity, internet-of-things, end-to-end connectivity, blockchain, big data, predictive analytics) for such endeavour is not an easy task. To some extent, this can potentially be traced back to the (over-)emphasis on "technology" when digital transformation topics are covered in the news and many academic papers. Often, technological innovation is in the focus and highlighted as the primary and at times only objective of *Digital Transformation* in the manufacturing industry and supply networks. However, the focus should be beyond technology adoption and aimed at developing "digital capabilities".

However, this "techno-centric" approach is doomed to fail in many cases if the underlying business processes are not critically evaluated and adapted to the changing (industrial) environment. Business processes should be properly (re-)engineered before being "digitalized" to exploit higher effectiveness levels enabled by the capabilities of the carefully selected digital technologies. Incorporating digital technologies such as smart sensor systems, 5G-networks, robotic process automation, cloud computing, data analytics, etc. requires rethinking the existing business process and envisioning the impact of their enabling digital technologies on a broader scale to capitalize on their digitalization opportunity.

In the following, this paper briefly reflects on the current state-of-the-art in the related topics of business process modelling and (re-)engineering and digital transformation, and the digital technologies and their implementation frameworks available to support successful business processes digitalization.

2.1 Business Processes & Digital Transformation

Digital Transformation refers to the increase of digitalized business processes in an enterprise or supply chain resulting in the adoption and integration of information, communication, and operational technologies to create new and enhanced digitally-enabled operations. Furthermore, the application of business process optimization and re-engineering techniques has been recognized as a prerequisite for successful business processes digitalization to offer a high degree of contextuality and specificity to the transforming enterprise or supply chain [7, 8].

However, even successful cases of proposed *Digital Transformation Frameworks,* for instance [8], struggle to formalize the operational levels of digital transformation. It seems that there is no clear distinction between (i) the digital transformation paths followed to digitalize an existing business process, and (ii) the engineering of a 'born-digital' business process based on newly available digital technologies.

2.2 Digital Technologies & Digital Transformation Frameworks

We have recently crossed the 10th anniversary of the Fourth Industrial Revolution, or Industry 4.0 era, which came alongside the introduction of new digital technologies such as "cyber-physical systems". These technologies enable us to connect the physical with the virtual world and bring forth tremendous (digital) opportunities as well as challenges for the design of next-generation production systems, specifically including their supply networks. The core principles of connectivity, virtualization, and data utilization [9] can take on various forms and provide numerous benefits across companies, supply networks, and even whole industries such as operations with higher visibility, transparency, predictability, and adaptability levels. Various innovative digital technologies are associated with this new industrial transformation, which is often wrongly reduced to the implementation of such technologies. As previously argued, it is believed that this "techno-centric" approach is posing a significant problem and it needs to be approached more strategically, using a combined approach of carefully reassessing the status quo of business processes and (digital) technologies, and then defining their future state more holistically.

2.2.1 Digital Technologies

Digital transformation efforts are often reduced to new digital technologies adoption, without including the pursuit of innovative and digitally-enabled business and operating models. Supporting this "techno-centric" approach, there is a large body of literature covering the characteristics and enabling factors of these new digital technologies. Prominent recent examples with significant citations include [10, 11, 12] among several others. While the various papers identify slightly different technologies and technology clusters associated with digitalization (and smartening) efforts, overall, these are consistent in what technologies are covered. A meta-analysis can extract 10 key digital technology clusters [13]:

(i) artificial intelligence, machine learning, and advanced simulation; (ii) cloud, fog, and edge computing; (iii) additive manufacturing, (iv) industrial internet-of-things and cyber-physical systems; (v) augmented reality, virtual reality, and digital twins; (vi) automation and robotics; (vii) cybersecurity; (viii) blockchain; (ix) smart sensor systems; and (x) 5G-networks.

All these technologies have received significant attention from industry, academia, and mainstream media over the last decade. Their business applications are evolving and the pressure for companies to at least have some active projects deploying some of these technologies is increasing. The tangible results however are sobering with many of these projects ending up with a 'successful' proof-of-concept, yet rarely transitioning into productive and value-adding use in enterprise-wide operations. Often termed the "pilot purgatory" [14], this challenge can partially be traced back to the lack of flexibility and awareness of the procedural requirements of the operations – and how they align with what advances the new digital technologies offer.

2.2.2 Digital Transformation Frameworks

Given the novelty and presence of new digital (and smart) manufacturing paradigms in trade publications, daily news, and overall public discourse it is not surprising that companies feel pressure to engage in activities targeting the adoption of digital technologies. To support this adoption, there are an increasing number of support systems available, ranging from maturity models [15] to dedicated decision support systems [16]. In the process, the manufacturing community also realized that support requirements for small and medium-sized enterprises differ from multinationals significantly [17]. Hence, dedicated models emerge, developed for various industries [18], company sizes [19], and application areas [20].

It has to be noted that most of the support systems are either targeting a generic, high-level, or a very defined process or technology. Rarely do these support systems address both comprehensively and allow for a critical assessment of current business processes and technological aspects at the same time. A reality that is very surprising to see in the industrial practice considering the academic efforts over the last decades of providing enterprise architectures and frameworks for business-IT alignment, and information systems integration and/or interoperability [21].

3. Research and Practice Gaps of Current Digital Transformation Approaches

The ongoing divide between the domains of business process optimization/(re-)engineering and technology implementation has been extensively studied in the existing scientific literature provided by the Enterprise Architecture discipline under many *Business-IT Alignment Frameworks* [21]. Moreover, this disconnection between technological capabilities (i.e., 'how to do it') and the expected business value of their enablement (i.e., 'should a particular (combination of) technology cluster(s) be applied') is frequently observed in the industrial practice – and regularly causes issues in many digital transformation efforts.

While business processes excellence and technology clusters excellence are often achieved individually, the connection and understanding of their interdependencies seem to be lacking, mainly in industrial practice. Significant research in business processes innovation regularly refers to classic literature [22]. However, the majority of publications on "business processes engineering" focus on the need to reinvent the business processes and only addresses technology as a subordinate topic. A possible reason is the comparative novelty of digital transformation efforts – with the rapid growth within the last decade – and the majority of business process literature emerging before that threshold seems to be one important aspect of this disconnection.

Although there are multiple alternatives to map business processes, no dominant one exists [23] that is universally accepted as a standard. Neither is explicitly considering the employed resources for a process task. While the *Business Process Modelling Notations (BPMN)* are not explicitly limited to specific processes, only the *Supply Chain Operations Reference (SCOR) model* [24] explicitly aims to cover the business processes beyond the individual organization. However, despite its targeted application to supply chain management, it does not go beyond observing the employed resources if at all, even in its revised edition [25]. The more recent *Digital Capabilities Model (DCM)* [26] considers the technological requirements and maps these onto two levels of detail for the business processes model, which includes the inter-connectivity of sub-processes. However, the business processes are not detailed to task flows, so it is not clear whether it can serve to leverage the (digital) technologies to optimize the business processes. In addition, neither of the described models covers events or results provided by external roles such as roles outside direct managerial control. Except for entirely, vertically integrated enterprises (from primary resources to final consumer), business process tasks depend on external inputs from suppliers, distributors, or others. Yet, technologies that would facilitate the interaction of roles across these supply chain partners are not mapped.

Approaches comparable to *Value Stream Mapping* [27], which is a methodology enabling the reduction of waste in manufacturing settings, do not seem to exist for evaluating the quality of a given business process. Applications in process mining [28] focus on the three aspects of discovery, conformance, and enhancement but do not investigate the process flow in terms of roles in swim lanes, employed technology, or data source – essential components of any digital transformation project.

The more financially focused *Cost-to-Serve Methodology* derives the cost of a given business process to deliver customer value [29]. The underlying assumption is that business processes excellence drives down cost-to-serve but given the lack of a process excellence measure, it remains uncertain whether any improvements have reached a global cost optimum.

4. Multiple Use Cases as Evidence for Common Pitfalls of Business Processes and IT Integration

Given the dearth of available research, the analysis of multiple case studies, derived and anonymized from recent consulting engagements, supports and highlights the discovery of typical shortfalls in business processes re-engineering in its practical application today. Across differences regarding industries and/or products, the case companies exhibit and report similar shortfalls across the board. The investigated domains in this paper to identify shortfalls transition from intra-organizational focus to inter-organizational, and include (i) decision engineering, (ii) intra-process quality, (iii) intra-company process alignment, (iv) inter-company process synchronization, and (v) process-technology alignment. To facilitate and enable a better comparison of the identified process shortfalls and control for some bias, only small and medium-sized enterprises in the manufacturing sector were included.

Table 1. Analyzed Case Studies

Use Case	**Core Product Segment**
Apparel	Fiber manufacturer
Home-appliance electronics	Handheld electronic devices
Medical supplies	Medical supplies and consumables
Elevator manufacturer	Elevators

- **Process & Customer Value Alignment:** Shortfalls in this domain result from a misalignment between customer value, which is achieved from decision-making and the executed business process. Colloquially, the process roles in such a shortfall are busy but not effective.
    - *Apparel* – a lot of data, no clarity of value generation. The case company had been collecting data from internal and external sources to understand the supply and demand patterns along its value chain. The results were synthesized in a digital twin of its value chain to prepare reports and dashboards. However, the key challenges were that the customer of such analysis was not involved in the design process, so the backwards-looking analysis was not able to provide value for forward-looking decision-making.
- **Process Completeness & RACI Consistency:** Even for cases in which a business process generates value; a shortfall may be the completeness and consistency of role assignment. The process only works from end-to-end because of exceptions, for instance, roles disregard the process to accomplish activities, or the activities are not recognized as dependent, and issues are resolved only during escalations.
    - *Home-appliance electronics* – lifecycle management decoupled from operations. For a home-appliance manufacturer, the business process had been mapped – also to a high degree ofgranularity following existing best practices – within functional boundaries. The challenge of missing inter-functional ("cross-silo") coordination was only noticed when manufacturing outsourcing initiatives stalled. The existing cross-functional gaps were no longer manageable within the organization and caused 'supplier-relationship' discord during new product introductions when sales plans could not align with supply plans because the bills-of-material were not released yet from engineering and the contract manufacturer did not receive the demand signal to prepare machine capacity. An issue that did not exist before when inter-personal relationships ensured that all functional departments were aware that new products were about to be released.
    - *Medical supplies* – a process making use of an advanced planning system does not address the required process coordination across functions. For a medical supplies manufacturer, the rollout of an advanced planning system in its supply chain function was expected to bring many improvements. When the expected improvements in supply responsiveness did not materialize, the root cause was eventually identified as ad-hoc plan reviews between the supply planners and the manufacturing lead. To improve overall plant performance, the plans

    were reviewed, manually improved and then modified in the system. However, the shortfall in supply responsiveness was unattainable in the first place due to gaps in the system configuration.

- **Process Alignment:** In particular, companies with multiple facilities/plants/locations, but also in cases where more than one employee owns a process role, are susceptible to this shortfall. This shortfall may exist in multiple processes for the same plants (process boundaries).
    - *Medical supplies* – organizational alignment to execute processes diverges between plants. At a manufacturer of medical supplies, its processes across plants even for the same product categories and with supposedly templated IT infrastructure were unable to standardize their intra-plant business processes to enable top-down management control for increased supply chain agility. The reasons were that individual legal structures resulted in slightly different reporting lines and stakeholders' management, the templated IT infrastructure had slight differences in the order fulfilment for tax reasons (e.g., under which customer order status inventory was consumed), when activities like production declaration occurred and were entered, and multiple others. Each amalgamation of reasons proved intractable without introducing additional systems support.
- **Process Synchronization beyond the Silo:** The inability to manage external business process partners directly may lead to shortfalls in synchronizing across departmental functions (i.e., "silos") and across supply network partners.
    - *Medical supplies* – a business process requires activities to be carried out by suppliers or customers, but this will not be modelled in a traditional business process map. For a medical supplies manufacturer, interactions with customers (i.e., medical care providers) were challenging. The order fulfilment is required to input in terms of inventory levels under consignment at the customer and their consumption, as well as confirmation of proposed replenishments. The timing of the input delivery was questionable and data quality was uncertain. The resulting poor quality of replenishments was compensated by customer reviews of the orders which in turn resulted in changes in quantities and potential delays in approving the replenishments. While these operational issues are not unusual, the business process mapping did not indicate the extent of challenges and the repercussions in other activities in the business process.
- **Process & Technology Integration:** The quality of the business process mapping and the technology choice can fall short in optimizing the process by rethinking the required activities and the employed technology.
    - *Elevator manufacturing* – process redesign ignores technology leverage. For an elevator manufacturer, the business process innovation of the order fulfilment was approached with management buy-in and bottom-up involvement to address shortcomings in the existing process. The challenge occurred when, after the future business process was fixed, the hand-over to the IT function indicated that several technology solutions were ignored (e.g., increased supply chain visibility in the planning systems, push notifications instead of periodic checks for order updates, switch to a new project management system to replace the ageing one, etc.). Thus, the business process was optimized within the given organizational hierarchy and systems landscape, but the process was not optimized.

The five identified types of shortfalls for turning business processes into "digital operations", seem to be recurring. While not an exhaustive study of industry cases, two cases per shortfall type are provided. Despite the occasionally multiple shortfalls per use case, it should be noted that most of these companies have been able to grow in terms of revenue and/or profit for several years. This indicates that the shortfalls are not operational blockers. However, it must be noted that these shortfalls are observable at the end of multi-year digital transformation programs. While the analyses do not detail the extent of the shortfalls at the beginning of these digitalization efforts it is safe to assume that the desire is to improve efficiency in the business operations (i.e., one or more business processes).

How these shortfalls are effectively bridged has not been exhaustively studied. However, in each industry case, the prevalence of manual intervention and the use of Microsoft Excel on the lower hierarchical levels is obvious. These companies have been visited regularly by one of the authors of this paper and the computer screens of the industry case employees rarely show integrated systems.

5. Discussion and Research Agenda

To support end-to-end supply networks' digitalization towards "digital operations", business processes and their IT must be well aligned. This implies as discussed that business processes and IT models of different supply network partners become integrated and/or made interoperable with each other. For such endeavour, three different approaches are commonly discussed in the literature [30]. The first approach focuses on the *IT alignment problem* and addresses it with *service-oriented architectures* to support information systems interoperability [31, 32], the second approach focuses on *business processes alignment* to make them executable across supply network partners by using *orchestration* or *choreography languages* [33], and the third approach focuses on *business-IT alignment methods* for detecting and correcting misalignments in a "two-way" [34]. These approaches have been well studied, nevertheless, these remain rather vague in the industry in terms of how to define and practice their "threefold alignment" [30] to sustain digital supply networks integration as business processes requirements and digital technologies evolve [35, 36].

Moreover, the fusion of business (processes) and IT strategies, their ideal alignment, in a Digital Economy (or Industry 4.0 era) seems to be inevitable. A phenomenon that nowadays practitioners and researchers refer to as "Digital Transformation", given the rising importance of digital technologies for the competitiveness of business processes. Therefore, companies and their supply networks need to aim beyond *business-IT alignment* in their business processes digitalization efforts and understand the differential value of digital technology for enhanced business processes performance, and how digital and hybrid business processes can leverage new digital capabilities such as visibility, transparency, predictability, and adaptability for the next-level of business processes performance and competitiveness [37].

Lastly, as the debate continues on how to successfully transform the operations of companies and their digital supply networks into high performance and competitive "digital operations", three indispensable research and practice lines emerge for the *Next Generation of Business-IT Alignment Frameworks,* also referred to as *Digital Transformation Frameworks* [38]: (i) How to improve the value derived from IT and digital technologies by using data analytics and machine learning solutions, (ii) How to enable agility by tapping into "cloud" scalability to increase or decrease IT resources as needed to meet the changing demand, and (iii) How to couple digital technologies with business strategies for the new "digital business strategies" required for a renewed Digital Economy.

6. Conclusions and Further Research

This paper has identified business processes as a potential resolution for the limited success in digital operations in companies and their supply networks. Scientific and grey literature, as well as multiple industry cases, indicate a research and practice gap in the successful digitalization of business operations; while the existing literature focuses mainly on the technological aspects of business processes digitalization efforts, the business processes (re-)engineering aspects appear to have been neglected in a much-needed business-IT alignment for successful digital transformations.

The two main limitations of this research paper are the selection of industrial use cases and the scope of the literature review. The industry cases may be a self-selected group, which has more process shortfalls than the population of cases. Also, the examined cases may not be an exhaustive selection of all possible process shortfalls. An extended scope of literature review, possibly including more grey literature, may yield more results on the aspects of business process quality assessments.

Further research shall include a more detailed guideline to resolve the identified shortfalls, either individually or holistically. To do so, a detailed analysis of the enablers for continued operations despite the shortfalls can be expected to be helpful.

7. References

[1] Alicke, K., Barriball, E., Trautwein, V., 2021. How COVID-19 is Reshaping Supply Chains, https://www.mckinsey.com/business-functions/operations/our-insights/how-covid-19-is-reshaping-supply-chains

[2] Schaller, R.R., 1997. Moore’s Law: Past, Present and Future. IEEE Spectrum, 34(6), 52–59, https://doi.org/ 10.1109/6.591665

[3] Hua, N., Buchta, H., Zheng, X., Zhang, H., Zhou, B., 2009. Performance Analysis of an Improved Postponed Lightpath Teardown Strategy in Multi-Layer Optical Networks. Proceedings of the 2009 Asia Communications and Photonics Conference and Exhibition, Shanghai, China, https://doi.org/ 10.1364/ACP.2009.WF3

[4] Gupta, P., Wildani, A., Miller, E.L., Rosenthal, D., Adams, I.F., Strong, C., Hospodor, A., 2014. An Economic Perspective of Disk vs. Flash Media in Archival Storage. Proceedings of the IEEE 22nd International Symposium on Modelling, Analysis & Simulation of Computer and Telecommunication Systems, Paris, France, https://doi.ieeecomputersociety.org/10.1109/MASCOTS.2014.39

[5] Tobola, J., Kořenek, J., 2011. Effective Hash-based IPv6 Longest Prefix Match. Proceedings of the 14th IEEE International Symposium on Design and Diagnostics of Electronic Circuits and Systems, Cottbus, Germany, https://doi.ieeecomputersociety.org/10.1109/DDECS.2011.5783105

[6] Cecere, L., 2021. Navigating The Supply Chain Through The Pandemic: Opportunity to Build Better. Research Report, Supply Chain Insights, Hanover, USA, https://content.kinaxis.com/snop/report-navigating-pandemic

[7] Fischer, M., Imgrund, F., Janiesch, C., Winkelmann, A., 2020. Strategy Archetypes for Digital Transformation: Defining Meta Objectives using Business Process Management. Information & Management, 57(5), 103262, https://doi.org/10.1016/j.im.2019.103262

[8] Romero, D., Flores, M., Herrera, M., Resendez, H., 2019. Five Management Pillars for Digital Transformation Integrating the Lean Thinking Philosophy. Proceedings of the 25th International ICE-Conference on Engineering, Technology and Innovation, Valbonne Sophia-Antipolis, France, https://doi.org/10.1109/ICE.2019.8792650

[9] Sinha, A., Bernardes, E., Calderon, R., Wuest, T., 2020. Digital Supply Networks: Transform Your Supply Chain and Gain Competitive Advantage with New Technology and Processes. McGraw-Hill Education.

[10] Franke, A.G., Dalenogare, L.S., Ayala, N.F., 2019. Industry 4.0 Technologies: Implementation Patterns in Manufacturing Companies. International Journal of Production Economics, 210, 15–26, https://doi.org/10.1016/j.ijpe.2019.01.004

[11] Mittal, S., Khan, M. A., Romero, D., Wuest, T., 2019. Smart Manufacturing: Characteristics, Technologies and Enabling Factors. Proceedings of the Institution of Mechanical Engineers, Part B, Journal of Engineering Manufacture, 233(5), 1342–1361, https://doi.org/10.1177/0954405417736547

[12] Ghobakhloo, M., 2019. Determinants of Information and Digital Technology Implementation for Smart Manufacturing, International Journal of Production Research, 58(8), 2384–2405, https://doi.org/10.1080/00207543.2019.1630775

[13] Wuest, T., Romero, D., Khan, M., Mittal, S., 2022. The Triple Bottom Line of Smart Manufacturing Technologies: An Economic, Environmental, and Social Perspective. pp. 310-330. In: Kurz, H. et al. (Eds.). Handbook of Smart Technologies: An Economic and Social Perspective. Routledge, London, https://www.taylorfrancis.com/chapters/edit/10.4324/ 9780429351921-20/

[14] Aramayo-Prudencio, A., Coxon, M., de Boer, Enno, De Ocampo, D., Kadocsa, A., Mühlreiter, B., van Niel, J., 2018. Digital Manufacturing – Escaping Pilot Purgatory. In: Behrendt, A. et al. (Eds.). Digital McKinsey, digital-manufacturing-escaping-pilot-purgatory.pdf (mckinsey.com).

[15] Schumacher, A., Erol, S., Sihn, W., 2016. A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises. Procedia CIRP, 52, 161–166, https://doi.org/10.1016/ j.procir.2016.07.040

[16] Liebrecht, C., Lang, M., Schaumann, S., Stricker, N., Wuest, T., Lanza, G., 2021. Decision Support for the Implementation of Industry 4.0 Methods – Toolbox, Assessment & Implementation Sequences for Industry 4.0. J. of Manufacturing Systems, 58, 412–430, https://doi.org/10.1016/ j.jmsy.2020.12.008

[17] Mittal, S., Khan, M., Romero, D., Wuest, T., 2018. A Critical Review of Smart Manufacturing & Industry 4.0 Maturity Models: Implications for Small and Medium-sized Enterprises (SMEs). Journal of Manufacturing Systems, 49, 194–214, https://doi.org/10.1016/j.jmsy.2018.10.005

[18] Rafael, L.D., Jaione, G.E., Cristina, L., Ibon, S.L., 2020. An Industry 4.0 Maturity Model for Machine Tool Companies. Technological Forecasting and Social Change, 159, 120203.

[19] Mittal, S., Khan, M., Purohit, K., Menon, K., Romero, D., Wuest, T., 2020. A Smart Manufacturing Adoption Framework for SMEs. International Journal of Production Research, 58(5), 1555–1573.

[20] Caiado, R.G.G., Scavarda, L.F., Gavito, L.O., Ivson, P., de Mattos Nascimento, et al., 2021. A Fuzzy Rule-based Industry 4.0 Maturity Model for Operations and Supply Chain Management. International Journal of Production Economics, 231, 107883, https://doi.org/10.1016/ j.techfore.2020.120203

[21] Zhang, M., Chen, H., Luo, A., 2018. A Systematic Review of Business-IT Alignment Research with Enterprise Architecture. IEEE Access, 6, 18933–18944, https://doi.org/10.1109/ACCESS.2018. 2819185

[22] Hammer, M., Champy, J., 1993. Reengineering the Corporation: A Manifesto for Business Revolution. Harper Business, 4th Edition.

[23] Wang, W., Ding, H., Dong, J., Ren, C., 2004. A Comparison of Business Process Modeling Methods, IBM, rc23988.pdf (ibm.com).

[24] Poluha, R., 2007. Application of the SCOR Model in Supply Chain Management. Illustrated Edition, Cambria Press.

[25] MMH Staff 2017. APICS to Update Industry-Recognized SCOR Model in its 20th Year. Supply Chain Management Review, https://www.scmr.com/article/apics_to_update_industry_recognized_scor_ model_in_its_20th_year

[26] Association for Supply Chain Management, 2022. Digital Capabilities Model for Supply Networks, https://dcm.ascm.org/

[27] Rother, M., Shook, J., 1999. Learning to See. Lean Enterprise Institute.

[28] Van Der Aalst, W., 2016. Process Mining: Data Science in Action. 2nd Edition, SpringerLink.

[29] Braithwaite, A., Samakh, E., 1998. The Cost-to-Serve Method. The International Journal of Logistics Management, 9, 1, 69–84, https://doi.org/10.1108/09574099810805753

[30] Kassahun, A., Tekinerdogan, B., 2020. BITA*: Business-IT Alignment Framework of Multiple Collaborating Organisations. Information and Software Technology, 127, 106345, https://doi.org/ 10.1016/j.infsof.2020.106345

[31] Chen, H.-.M., 2008. Towards Service Engineering: Service Orientation and Business-IT Alignment. Proceedings of the 41st Annual Hawaii International Conference on System Sciences, https://doi.org/ 10.1109/HICSS.2008.462

[32] Cuenca, L., Boza, A., Ortiz, A., J.M. Trienekens, J., 2014. Business-IT Alignment and Service Oriented Architecture – A Proposal of a Service-Oriented Strategic Alignment Model. Proceedings of the 16th International Conference on Enterprise Information Systems, https://doi.org/10.5220/0004973204900495

[33] Karande, A., Karande, M., Meshram, B.B., 2011. Choreography and Orchestration using Business Process Execution Language for SOA with Webservices. International Journal of Computer Science Issues, 8(2), 224–232, https://ijcsi.org/articles/

[34] Chen, H.-.M., Kazman, R., Garg, A., 2005. BITAM: An Engineering-principled Method for Managing Misalignments between Business and IT Architectures. Science of Computer Programming, 57(1), 5–26, https://dl.acm.org/doi/10.1016/j.scico.2004.10.002

[35] Plomp, M.G.A., Batenburg, R.S:, 2014. Interorganisational Information Systems Maturity: Do Supply Chain Integration and Business/IT-Alignment Coincide?. BLED e-Conference Proceedings. 41. http://aisel.aisnet.org/bled2014/41

[36] Maes, R., Rijsenbrij, D., Truijens, O., Goedvolk, H., 2000. Redefining Business-IT Alignment through a Unified Framework. PrimaVera Working Paper 2000-19, Amsterdam: University of Amsterdam, https://hdl.handle.net/11245/1.182103

[37] Kahre, C., Hoffman, D., Ahlemann, F., 2017. Beyond Business-IT Alignment – Digital Business Strategies as a Paradigmatic Shift: A Review and Research Agenda. Proceedings of the 50th Hawaii International Conference on System Sciences, https://aisel.aisnet.org/hicss-50/os/digital_innovation/2/

[38] Jonathan, G.M., Rusu, L., Wim, V.G., 2021. Business-IT Alignment and Digital Transformation: Setting A Research Agenda. Proceedings of Information Systems Development (ISD2021), https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1624551

8. Biographies

Dr. Andreas M. Radke is a Senior Executive Consultant at mSE Solutions since 2019 and is based in the Chicago office (since 2020). His work focuses on end-to-end supply chain digitalization, business process optimization, and digital supply network collaboration. He previously worked for Bayer AG (2018-2019) and other companies. He was CIRP Research Affiliate from 2012-2014 and received his Ph.D. from the Hong Kong University of Science & Technology in 2012. He was BVL Chairperson of the BVL Chapter Singapore from 2017 to 2020.

Dr. Thorsten Wuest is an Associate Professor of Smart Manufacturing at the Industrial and Management Systems Engineering Department of West Virginia University (since 2015). He has worked as a research scientist for BIBA – Bremer Institut für Produktion und Logistik GmbH in Bremen since 2009. He was named inaugural J. Wayne and Kathy Richards Faculty Fellow in Engineering in 2017.

Dr. David Romero is a Professor of Advanced Manufacturing at the Industrial Engineering Department of the Tecnológico de Monterrey (university). He is an Agenda Contributor at the World Economic Forum Council on "Shaping the Future of Advanced Manufacturing and Value Chains", and a World Manufacturing Forum Ambassador. He serves at different editorial and scientific committees and advisory boards in academia, industry, and government in the disciplines of business and industrial engineering. Nowadays is focused on promoting the concepts of the "Operator 4.0" and "Digital Lean Manufacturing"; father of both terms.

3rd Conference on Production Systems and Logistics

Additive Manufacturing Production Shops: A Requirements Analysis

Günther Schuh[1], Gerret Lukas[1], Steffen Hohenstein[1], Viktoria Krömer[1], Niklas Lindholm[1]
[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany*

Abstract

Additive Manufacturing (AM) technologies become increasingly relevant for manufacturing companies. Despite having the highest share of AM applications, end-use parts are mostly used for spare or special parts and rarely within series applications. This paper addresses the challenge of practically implementing AM series production into industrial environments by means of a requirements analysis. It proposes a methodology on how to record, prioritize and meet requirements for AM production shops.

Successful implementation demands understanding of the requirements for AM production shops from both a factory and an AM perspective. Quality Function Deployment (QFD) is chosen as a methodology for the requirements analysis. It offers a framework for structured collection and weighting of the requirements identified through expert interviews with AM users and system manufacturers. Subsequently, measures and a basic plan of action on how to implement AM series production into production shops are defined.

The analysis reveals seven requirements for AM production shops within the categories spatial organization, process chains and flow systems. Most of them concern process chains, making these primary obstacles towards additive series production on the technical side. Substantial requirements are high process stability, fast process chains and the reduction of manual post-processing.
Different advancements are necessary on the AM and the factory side. On the factory side, measures that form synergies to conventional manufacturing technologies, such as cross-usable quality assurance systems, are favorable. On the AM side, focus lies on the enhancement of physical and digital process chains.
The results show that implementing AM production shops requires joint and interdisciplinary developments by AM users and system manufacturers. Further research and a larger sample are needed for validation as well as practical realization and advancement of the identified measures.

Keywords

Additive Manufacturing; Additive Manufacturing Production Shops; Additive Factory Structures; Quality Function Deployment; Requirements Analysis

1. Introduction

Conventional manufacturing technologies have reached full process maturity. They struggle to address the complexity of global market structures and customer requirements on individualization. To fulfill customer requirements and remain competitive, the flexibility of production systems becomes critical to success for manufacturing companies. These demands necessitate technological development and a redesign of production shops. [1]
In this context, additive manufacturing (AM) becomes increasingly relevant for manufacturing companies. AM technologies facilitate cost- and time-efficient as well as tool-free part production allowing for individualization and complexity for free. Profound development of AM technologies has led to a shift from prototyping and special part applications to industrial manufacturing of end-use parts [2]. Despite having the

DOI: https://doi.org/10.15488/12167

highest share of AM-applications by now, end-use parts are still rarely produced within series applications [3,4].

Therefore, this paper addresses the reasons preventing the implementation of AM series production in industrial environments by means of a requirements analysis for AM production shops.

2. State of the Art

This chapter gives an overview on industrial AM and applications. Further, research on the implementation of AM series production in production shops is introduced.

2.1 Industrial AM

AM technologies produce parts from 3D model data by joining material layer by layer, as opposed to subtractive and forming manufacturing technologies [5]. The technologies are classified in seven process categories according to DIN EN ISO/ASTM 52900 [6]. Additional to the mere manufacturing process, AM requires pre- and post-processing operations, inter alia for meeting part requirements, resulting in AM process chains. Though differentiation depending on the AM technology, Figure 1 displays a generic AM process chain.

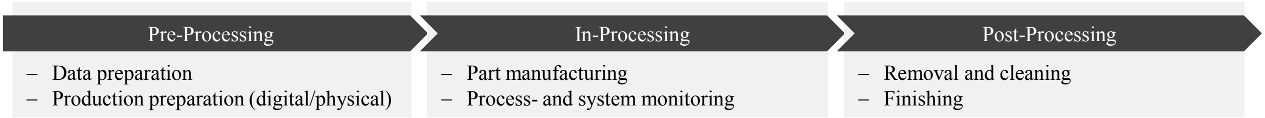


Figure 1: Generic AM process chain based on [7]

Industrial AM describes process chains with a maturity level high enough to compete with conventional manufacturing processes [8]. As of today, AM finds industrial use in highly complex and variable product programs or small batch sizes, for example in mass customization and mass complexity manufacturing applications. Further, AM facilitates strategies such as spare parts on-demand and digital warehouse. Additionally to these end-use part applications, low-volume series parts also begin to establish. The degree of the industrial integration of AM significantly depends on both industry and use-case [9].

2.2 Additive Series Production in Production Shops

According to HALEEM AND JAVAID, large-scale integration of AM series production in production shops necessitates the consideration of AM as a digitized manufacturing technology and the fulfillment of the principles of modern, networked factories [4]. BREUNINGER ET AL. identify needs for adaption in the following fields of a production system for the implementation of AM series production: Spatial organization, product design methodology, quality, organization of logistics, handling of material and process flow [10]. YI ET AL. address the integration of AM into manufacturing systems in the form of a holistic enterprise approach, emphasizing the need for quality assurance (QS) systems [11]. The main obstacles identified are a lack of know-how and high risks associated with the introduction of not yet widely established technologies. According to the authors, further research is needed regarding process chain and manufacturing-related barriers as well as counteractions. KYNAST ET AL. identify the use of automated process chains planned over the entire manufacturing process as a requirement on a continuous additive process chain integrated into a manufacturing system [12].

DERADJAT AND MINSHALL record requirements for mass customization within the segments technology, operations, organization and internal and external influencing factors through case studies. Identified requirements are, inter alia, front-end software solutions, simplification of material handling, speed and stability of process chains, part design, employee trainings as well as supplier chains and business models. [13]

2.3 Interim Conclusion

The state of the art shows the relevance of AM for series production. Requirements for the implementation of AM series production already exist. However, approaches lack a methodology for the assessment of these regarding priority, applicability, usability and specification. This causes a gap between the weighting of the requirements and the defined targets, making the implementation of AM production shops more difficult. Previous research does not holistically consider the role of factories as production systems in fulfilling requirements, but rather focuses on the optimization of mere AM processes. This research approach closes the gap by using Quality Function Deployment (QFD) to record, prioritize and meet requirements for AM production shops from both an AM and factory perspective.

3. Methodology

Following, the pursued methodology is introduced through the QFD, the application of a house of quality (HoQ) as well as the data collection.

3.1 Quality Function Deployment

QFD is a methodology developed for planning products and processes. It focuses on the voice of the customer as the foundation for product design. First, customer requirements on the usability of a product are recorded and considered in the product development process. Second, based on customer requirements, design specifications and competitive analyses are carried out. Additionally to the design and optimization of products, QFD represents a methodology for planning tasks along the entire company value chain. QFD facilitates the accomplishment of planning projects, such as strategy, organization and technology planning. [14] The results determined through QFD can be illustrated through a HoQ.

Building on the results, measures for the optimization and development of products and planning processes can be derived by applying QFD, making it a favorable methodology to conduct a requirements analysis for AM production shops.

3.2 Application of a House of Quality

The general setup of the HoQ is shown in Figure 2 and follows the subsequent steps [15,16]. Depending on the planning object, not all steps must be performed. In this research approach, the planning object consists of the integration of AM series production in a production shop. Two HoQ are built due to the separate consideration of measures on the AM system manufacturer and user side.

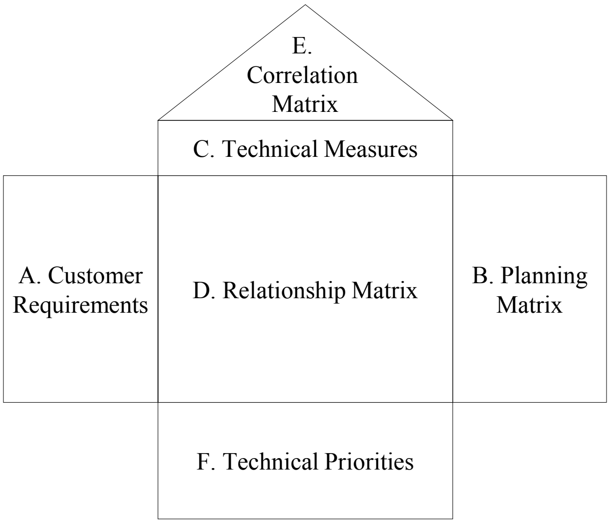


Figure 2: General setup of the HoQ

A. Customer requirements

Firstly, requirements that must be met to enable the use of AM in series production are recorded. Data collection tools serve this purpose. Secondly, the requirements are analyzed and interpreted to avoid doubling. Thirdly, sorting of the requirements into groups defined by the topics of the planning object takes place. Lastly, weighting of the requirements selected for the HoQ is conducted, taking into consideration additional data from market analyses.

B. Planning matrix

The planning matrix contains different representations, depending on the planning object. In the context of this research proposal, it reflects whether requirements must be implemented on the AM system manufacturer or on the factory side.

C. Technical measures

To fulfill the customer requirements recorded, technical measures are defined which are directly related to the requirements. Further, each measure requires measurable development potential. A target value and an optimization direction are assigned to each measure.

D. Relationship matrix

The relationship matrix shows the extent to which the technical measures contribute to meeting the identified customer requirements. Data from market analyses support this estimation. Predefined symbols represent the relationship between technical measures and customer requirements.

E. Correlation matrix

The correlation matrix, representing the roof of the HoQ, shows how the change of one technical measure affects other measures. Negative influences are considered separately in this step and taken into account during the planning phase. They represent the basis for trade-off decisions regarding the implementation of measures.

F. Technical matrix

In the final step of the QFD, a competitive comparison is carried out between the measures. For this purpose, market analysis data is evaluated with regard to target fulfillment of the measures. Following, their target fulfillment level is weighted against competing products.

3.3 Data Collection

Additionally to systematic literature research, expert interviews are conducted for the recording of requirements for the QFD. These are conceptualized on the semi-structured interview method, facilitating the obtainment of exclusive expert knowledge in the field of AM series production in industrial environments [17]. To integrate perspectives from different industrial sectors and companies into the requirements analysis, respectively four experts of both AM system users and system manufacturers are selected. To ensure optimum contribution to the research proposal, all interview partners have direct relation to either additive series or final part production.

A structured expert survey is conducted through a questionnaire in an online survey tool. Data collected through this method are used to quantitatively evaluate the aspects relevant to the research question. The design of the questionnaire is based on the elements of the HoQ with aspects to be collected mainly relating to the planning and relationship matrix. For measureable and comparable results, the questionnaire consists of only closed questions [18].

The majority of the questionnaire is based on an ordinal scale, describing the contribution of technological advancements to the fulfillment of requirements for AM series production. As a common instrument to determine the position of ordinal scales, the median of the results is used for evaluation. For a congruent data basis, questions with an interval scale are also evaluated using the median. The sample size is N=22, which consists of industrial AM users (27.3%), AM users in research institutions (36.4%), AM equipment manufacturers (22.7%) and other AM users and experts (13.6%). Most participants are AM department and project managers, as well as employees in research and development.

4. Results

The presentation of the results is structured according to the HoQ, including customer requirements, planning matrix, technical measures, relationships, correlations as well as a holistic integration.

4.1 Customer Requirements

As the aim of this study is the design of development measures for the integration of AM series production, only requirements relating to areas with further development needs are considered. All parties directly involved in the concept implementation for AM production shops are considered stakeholders of the planning object. The focus is on AM system users on the production shop and AM equipment manufacturers on the technological side. The identified requirements are weighted on a scale from “1” to “5” according to the results from the expert survey, with “1” having lowest and “5” having highest priority. Table 1 visualizes the recorded requirements, their classification according to structural factory areas as well as their weighting.

Table 1: Requirements identified through expert interviews

Structural Area	**#**	**Requirement**	**Definition**	**Weighting**
Spatial organization	i	Good integrability into existing processes	Adaption of manufacturing and post-processing operations of AM and conventional manufacturing	3
Process chains	ii	Fast process chains	Minimization of lead times	4
	iii	High process stability	Avoidance of unplanned downtimes and repeatable part quality	5
	iv	Little manual post-processing effort	Part separation from build plate, removal of support material	4
Flow systems	v	Simple material handling	Easy material supply and avoidance of material loss & contamination	2
	vi	Automated material flow	Integration of AM materials into an automated in-plant material flow system	2
	vii	Continuous information flow	Avoidance of information losses and media discontinuities within stations	2

4.2 Planning Matrix

Based on the requirements recording, the planning matrix states whether the responsibility for addressing the requirement is assigned to AM users (production shop) or AM system manufacturers (AM processes). According to the results, process chain related requirements must be addressed through AM processes. In particular, "fast process chains", "high process stability" and "little manual post-processing effort" are assigned to the AM process side by the survey participants. Flow system related requirements however are more likely to be addressed within the production shop. Material handling is identified as an important requirement. The fulfillment of simple material handling has a weak tendency towards implementation on the production shop side. The described tendencies remain mostly unchanged even under separate evaluation of the results from AM system manufacturers and users.

4.3 Technical Measures

In order to address the fulfillment of the requirements, measures for further developments on the AM system side (Table 2) and production shop (Table 3) are derived, dividing the analysis. For each defined measure on either side, a matching measure on the other side is defined to facilitate a comparison of both analyses. The definition of a target value and optimization direction enables the evaluation of the development status of each measure. These values are based on the results of the expert interviews and directly related to the previously defined requirements.

Table 2: Measures on the AM system side

#	Measure	Definition	Target	Optimization Direction
1	Communication capability of manufacturing equipment	Increasing degree of automation through interfaces enabling intelligent networking	Versatile interfaces	Improve
2	Software solutions of manufacturers	Software solutions to be integrated into PPS system	Cross-process planning and control	Fix target value
3	Machine robustness	Shielding of workspace against external influences (humidity, contamination)	No external influences	Fix target value
4	Integrated post-processing	Include post-processing steps, e.g. part cleaning or support removal into machine	Comprehensive	Improve
5	Material supply to machine	Improvement of interfaces for material supply to and material removal from machine	Closed materials cycle	Fix target value
6	In-situ quality control	Detection of defects during manufacturing process	Real-time analysis of production data	Fix target value

Table 3: Measures on the production shop

#	Measure	Definition	Target	Optimization Direction
7	IT infrastructure communication capability	Adaption of ERP system to simplify the integration of new	Versatile interfaces	Improve

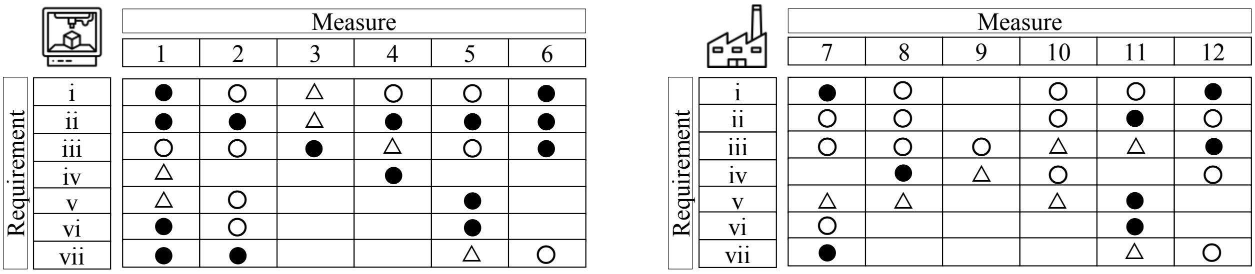
		technologies and resulting planning tasks		
8	Product design software	Tools to optimize product design for AM processes	AI optimization of manufacturing data	Improve
9	Machine environment	Avoiding temperature and humidity fluctuations as well as contamination	Clean room	Fix target value
10	Peripheral post-processing equipment	Camera systems to automatically detect and initiate necessary post-processing steps on assigned machines	Comprehensive	Improve
11	Material supply	Automated material flow on the production shop, e.g. through automated guided vehicles	Automated	Improve
12	Post production quality control	Automated quality control using non-destructive methods to detect e.g. surface finish, dimensional accuracy	Line-integrated	Fix target value

4.4 Relationships

The relationship matrix represents the body of the HoQ. Therefore, it is an important step within the final evaluation of the QFD. Input data for this step are derived from the results of the questionnaire. Participants answer the question, “How does an improvement of the following design fields contribute to the fulfillment of requirement X?” The response options are “not”, “weak”, “medium” and “strong”, resulting in an ordinal scale. Measures assessed as non-contributing are not displayed in the correlation matrix. Measures influencing the fulfillment of requirements are marked with symbols based on a common representation of the HoQ:

- Weak correlation: △
- Medium correlation: ○
- Strong correlation: ●

The implementation of measures with strong correlations is recommended for prioritization to meet requirements. Figure 3 shows the relationship matrix between requirements and measures for AM processes and production shops.



	Measure					
Requirement	1	2	3	4	5	6
i	●	○	△	○	○	●
ii	●	●	△	●	●	●
iii	○	○	●	△	○	●
iv	△			●		
v	△	○			●	
vi	●	○			●	
vii	●	●			△	○

	Measure					
Requirement	7	8	9	10	11	12
i	●	○		○	○	●
ii	○	○		○	●	○
iii	○	○	○	△	△	●
iv		●	△	○		○
v	△	△		△	●	
vi	○				●	
vii	●				△	○

Figure 3: Relationship matrix for AM processes (left) and production shops (right)

4.5 Correlations

The correlation matrix shows how the measures influence the fulfillment of each other, both positively (+) and negatively (-). It is developed based on data of today's industrial manufacturing processes and findings from the expert interviews. The correlation matrix for AM processes and production shops according to the defined measures in Table 2 and Table 3 are visualized in Figure 4.

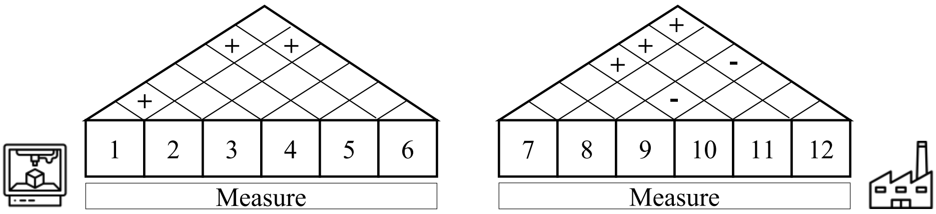


Figure 4: Correlation matrix between measures for AM processes (left) and production shops (right)

4.6 Integration of the Results

For integration and analysis of the results, the absolute weighting of the measures is determined. For this purpose, the symbols of the relationship matrix are assigned numerical values:

- Weak correlation "△" – 1
- Medium correlation "○" – 3
- Strong correlation: "●" – 5

To determine the total weighting of a measure and therefore the prioritization for implementation, the respective entry of the relationship matrix is multiplied by a requirement's weighting. The results are added up to the total weighting for each measure. Following formula is used for calculation:

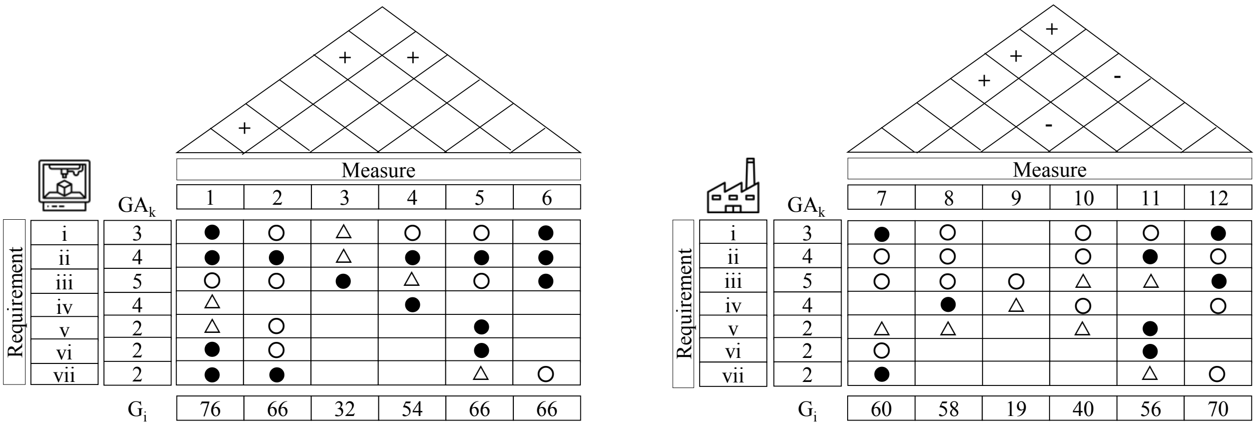
$$G_i = \sum_k GA_k * B_{ki} \tag{1}$$

G_i = Absolute weighting of measure i

GA_k = Weighting of requirement k

B_{ki} = Relationship matrix entry

The results are displayed in Figure 5.



		Measure					
	GA$_k$	1	2	3	4	5	6
i	3	●	○	△	○	○	●
ii	4	●	●	△	●	●	●
iii	5	○	○	●	△	○	●
iv	4	△			●		
v	2	△	○			●	
vi	2	●	○			●	
vii	2	●	●			△	○
	G$_i$	76	66	32	54	66	66

		Measure					
	GA$_k$	7	8	9	10	11	12
i	3	●	○		○	○	●
ii	4	○	○		○	●	○
iii	5	○	○	○	△	△	●
iv	4		●	△	○		○
v	2	△	△		△	●	
vi	2	○				●	
vii	2	●				△	○
	G$_i$	60	58	19	40	56	70

Figure 5: HoQ for measures in AM processes (left) and production shops (right)

5. Discussion and Conclusion

The evaluation of the requirements analysis reveals that requirements for an AM production shop must be met primarily through the development of process chains. The greatest need for further development is seen in the requirement for high process stability. Requirements for fast additive process chains and little manual post-processing effort also have high priority for the industrial use of AM. On the technical side, the development status of process chains within a production shop represents the greatest obstacle towards additive series production. Lower development need is attributed to requirements relating to flow systems within a factory.

This analysis allows drawing of initial conclusions on the greatest challenges in implementing an AM production shop at the current state. However, the derivation of further technological developments requires the consideration of sectors and technologies influencing the fulfillment of the requirements. Therefore, measures are evaluated revealing that further developments must be designed differently on an AM system manufacturer's perspective (AM process side) and AM system user's perspective (AM production shop).

Analyzing the weighting of the requirements, measures on the AM process side are mostly weighted higher than those on the AM system user side. The highest weighted measure is improvement of the communication capability of manufacturing equipment (measure 1). On the production shop side, measures that form synergies to conventional manufacturing technologies, such as cross-usable quality assurance systems, are favorable. Non-destructive quality assurance of every part is critical to additive series production, but can also be used for conventionally manufactured parts. On the AM process side, the main focus lies on the enhancement of digital process chains. Another focus area is the automation of physical processes. Additionally to the fully automated printing process, development activities by AM system manufacturers should lead to automation of up- and downstream process steps.

The results allow the identification of focus areas for further technological development for AM system manufacturers and users. They show that implementing AM production shops requires joint and interdisciplinary developments by AM system manufacturers and users. However, due to qualitative data collection, subjective bias of the results may occur. Therefore, the results must be examined with regard to the qualitative quality criteria of intersubjectivity and indication of the research approach. Further research and a larger sample are needed for validation as well as practical realization and advancement of the identified measures.

References

[1] Westkämper, E., Spath, D., Constantinescu, C., Lentes, J., 2013. Digitale Produktion, 1ed. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 75.

[2] Gebhardt, A., Kessler, J., Thurn, L., 2016. 3D-Drucken: Grundlagen und Anwendungen des Additive Manufacturing (AM), 2 ed. Hanser, München.

[3] Ampower, 2021. AMPOWER Report 2021: Metal Additive Manufacturing Degree of Industralization. https://additive-manufacturing-report.com/report-2021/metal-am-industrialization-2021/#003. Accessed 1 February 2022.

[4] Haleem, A., Javaid, M., 2019. Additive Manufacturing Applications in Industry 4.0: A Review. Journal of Industrial Integration and Management 04 (04).

[5] VDI-Gesellschaft Produktion und Logistik, 2014. Additive Fertigungsverfahren - Grundlagen, Begriffe, Verfahrensbeschreibungen.

[6] Deutsches Institut für Normung, 2018. DIN EN ISO/ASTM DIS 52900:2018: Additive Fertigung - Grundlagen - Terminologie.

[7] Gibson, I., Rosen, D., Stucker, B., 2015. Additive Manufacturing Technologies. Springer New York, New York, 1-18.

[8] Eyers, D.R., Potter, A.T., 2017. Industrial Additive Manufacturing: A manufacturing systems perspective. Computers in Industry 92-93, 208–218.

[9] Wohlers Asscoiates (Ed.), 2019. Wohlers Report 3D Printing and Additive Manufacturing State of the Industry, 1 ed, pp. 32.
[10] Breuninger, J., Becker, R., Wolf, A., Rommel, S., Verl, A., 2013. Generative Fertigung mit Kunststoffen, 1 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 23-112.
[11] Yi, L., Gläßner, C., Aurich, J.C., 2019. How to integrate additive manufacturing technologies into manufacturing systems successfully: A perspective from the commercial vehicle industry. Journal of Manufacturing Systems 53 (53), 195–211.
[12] Kynast, M., Witt, G., Eichmann, M. (Eds.), 2017. Anforderungen an integrierte Prozessketten in der Additiven Fertigung, 1 ed. Hanser, München, 16-22.
[13] Deradjat, D., Minshall, T., 2015. Implementation of additive manufacturing technologies for mass customisation, in: 24th International Conference of the International Association for Management of Technology (IAMOT 2015). International Association for Management of Technology, 2079-2094.
[14] Schubert, M.A., 1989. Quality function deployment-a comprehensive tool for planning and development, in: Proceedings of the IEEE National Aerospace and Electronics Conference. IEEE, pp. 1498–1503.
[15] Chan, L.-K., Wu, M.-L., 2002. Quality Function Deployment: A Comprehensive Review of Its Concepts and Methods. Quality Engineering 15 (1), 23-35.
[16] Hauser, J., Clausing, D., 1988. The House of Quality. Harvard Business Review.
[17] Bogner, A., Littig, B., Menz, W., 2002. Das Experteninterview, 1 ed. VS Verlag für Sozialwissenschaften, Wiesbaden.
[18] Reinders, H., Ditton, H., Gräsel, C., Gniewosz, B. (Eds.), 2011. Empirische Bildungsforschung: Strukturen und Methoden, 1 ed. VS Verlag für Sozialwissenschaften / Springer Fachmedien Wiesbaden, Wiesbaden, pp. 53.

Biography

Günther Schuh (*1958) is head of the Chair of Production Engineering at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh is member of the Board of Directors at the Fraunhofer Institute for Production Technology IPT in Aachen and director of the Research Institute for Rationalization e.V. (FIR) at RWTH Aachen University.

Gerret Lukas (*1991) holds a master's degree in industrial engineering and has been head of the department Corporate Development at the WZL since 2020.

Steffen Hohenstein (*1991) holds a master's degree in production engineering and has been head of the group Additive Manufacturing Systems at the WZL since 2020.

Viktoria Krömer (*1995) holds a master's degree in industrial engineering and has been a research associate in the group Additive Manufacturing Systems at the WZL since 2021.

Niklas Lindholm (*1997) is an industrial engineering student at RWTH Aachen University and student assistant in the group Additive Manufacturing Systems at the WZL.

3rd Conference on Production Systems and Logistics

Levels Of Autonomy In Production Logistics: Terminology And Framework

Michaela Krä[1], Leonard Eckart[1], Marinella Rito[1] and Johannes Schilp[1, 2]

[1] *Chair of Digital Manufacturing, Faculty of Applied Computer Science, Augsburg University, Am Technologiezentrum 8, 86159 Augsburg, Germany*

[2] *Fraunhofer Research Institution for Casting, Composite and Processing Technology – IGCV, Am Technologiezentrum 10, 86159 Augsburg, Germany*

Abstract

The increasing demand of flexibility in production systems influences the organisation of production logistics and enhances the role of autonomous resources for logistic tasks. In the current state of the art, there exists neither a common definition of the term "autonomy" in the production logistics context nor a generalised approach regarding the classification of autonomous resources depending on their characteristics as well as their skills. Due to this lack, difficulties appear when intending to integrate autonomous resources - that are implemented for logistic tasks - in the superior production control processes which aim to meet the key performance indicators of the production system.

This paper analyses in a first step the current use of terminology regarding autonomy and related terms like automation and self-x approaches in production logistics. Based on these results, a definition of "autonomy" for production logistics and a universal framework for classifying autonomous resources regarding their level of autonomy can be proposed. This allows to specify afterwards the appropriate level of autonomy in production logistics for a specific production system.

Keywords

Autonomy; Autonomous Transport System; Autonomous Guided Vehicle; Production Logistics; Production Control

1. Introduction

Globalization forces production companies to deal with high market dynamics, shorter product life cycles, increased competition, and rising volatility. Therefore, production systems need to cope among other challenges more and more with the customer demand of getting individualized products. This expectation leads to rising complexity and dynamics in production environments as well as production processes due to the necessary flexibility [1,2].

The current developments in the context of industry 4.0 concerning data exchange and interconnectivity in production systems offer various possibilities to analyse workflows in a more detailed way [3]. It is now possible to understand processes and their interdependences on different levels based on collected data and to hereupon optimize diverse parameters and target values, e.g. throughput time and/or product output [4,5]. In addition, this also highlights the significance of non-value-adding processes in production like production logistics as well as their importance for reaching key performance indicators (KPI) and emphasizes the importance of integrating them in communication and exchange processes [6,1]. In the context of logistics,

DOI: https://doi.org/10.15488/12140

this development is called “logistics 4.0” [7] and underlines that it is not preferable to look at intralogistics processes in an isolated way due to its influence on meeting planned production schedules and due dates [8]. Challenges for planning and control in this context consist in finding a connection between central and decentral approaches [9] and in integrating autonomous and intelligent systems [10]. Especially choosing an appropriate autonomous system is difficult for decision-makers as there is a lack of term definition and classification of levels of autonomy within the scope of production logistics.

With increasing dynamics, flexibility, and complexity as detailed above, an increasing decentralized and autonomous based organization of production logistics systems is required [7]. An exclusively central approach in production logistics is not sufficient because of the unpredictable environment. Therefore, decentralized approaches have to be taken into account [4] and conventional planning and control methods for logistic processes are no longer sufficient [11]. That’s why this paper introduces an appropriate definition of the term “autonomy” for production logistics based on an analysis of the current use of the term in state of the art publications (cf. chapter 3). Afterwards, a universal framework for classifying autonomous systems regarding their level of autonomy is described (cf. chapter 4). The developed framework supports decision-makers in manufacturing companies to choose a proper autonomous transportation system with relevant characteristics referring to a corresponding application.

2. State of the art

In this section, basic principles of production logistics and applied resources are presented to frame the analysis as well as the developed definition and the framework explained afterwards in chapter 3 and 4. The whole topic has a non-neglectable connection to production planning and control processes. So, they are briefly introduced in the beginning.

2.1 Production logistics

Within a production organization, production logistics deal with the planning and control of material and information flow. In this context, production logistics are placed between procurement logistics and distribution logistics and comprise all activities to supply production and assembly processes with material (raw material, operation material, semi-finished goods or purchased goods) as well as the transportation of semi-finished or finished products to the next production step or the stock [12]. The main goal of production logistics is the on-time delivery of material on the one hand to avoid costs for downtimes due to delays and on the other hand to prevent high waiting times in case of too early deliveries [13,14]. So, there exists an important influence on throughput time [8]. Current challenges in production logistics are induced by the changes due to industry 4.0 approaches and comprise especially ensuring the logistic flow in uncertain and changing production environments as well as the integration in higher level control processes [6,15].

2.2 Production planning and control

The main tasks of production planning and control in manufacturing systems are generating a valid production program based on orders, task allocation and production supervision in order to reach logistic KPIs [16,17]. Planning and control is introduced here briefly because of the interaction and relation between the superior planning and control level and the executing logistic level: a transport system is not able to operate without respecting other processes in the manufacturing system and impacts overall KPIs. Basic logistic KPIs in production are for instance throughput time (time between order approval and order completion), inventories (amount of orders that are approved but not yet completed), utilization (ratio between average output and maximum output of a production resource or system) and delivery reliability (amount of orders that are completed within the planned delivery time) [19,18,23,20,22,21]. Logistic KPIs that are relevant in the context of production logistics are in general derived based on customer needs - here the requirements of value-adding manufacturing processes - and therefore include objectives as delivery time, delivery lateness, and delivery reliability [18,21]. For more detailed information on planning and control see for example [18,24,16,17].

2.3 Autonomous transport system

In this paper an autonomous transport system (ATS) is defined as a fleet of autonomous vehicles. The terms autonomous guided vehicle (AGV) and autonomous vehicle are used synonymously and describe vehicles without a driver that fulfil transport tasks in production logistics. Depending on the manufacturer and respectively the model, they can have differing skills and competences in order to complete transportation tasks. "Modern" shopfloor layouts and flexible organisation processes require intelligence on transport resource level to reach adaptability. More detailed information can be for instance found in [21,25,26].

3. Analysis regarding the use of the term "autonomy" in production logistics

The goal of this chapter is to derive a definition of the term "autonomy" in context of production logistics. Therefore, an analysis of the current use of terminology regarding autonomy and related topics is required.

3.1 Comparison autonomy – automation – self-x-approaches

Within a literature review, the main terminology differentiation between the terms autonomy, automation, and self-x is demonstrated in this subchapter. Subsequently, all central ideas are summarized and compared regarding abilities of considered system resources. Relevant literature is listed in Table 1 subdivided by their focus regarding differentiation of terminology.

Table 1: Classification of literature in context of production systems

Authors	**Autonomy**	**Autonomy and automation**	**Autonomy and self-x**
Windt et al., 2008 [27]	X		
Dumitrescu et al., 2018 [28]	X		
Gamer et al., 2019 [29]		X	
Müller et al., 2021 [30]		X	X
Stock et al., 2020 [31]			X
Scholz-Reiter and Höhns, 2006 [20]			X
Schuhmacher and Hummel, 2020 [22]			X

[27] describe the term autonomy in the context of autonomous control by processes with decentralized decision-making and the ability of system elements to make decisions independently. Furthermore, the authors characterize autonomous control in logistic systems "by the ability of logistics objects to process information, to render and to execute decisions on their own". The superior goal of the autonomous control is the increase of system robustness of non-deterministic system behavior and positive emergence through objective achievement of every single logistics object. Accordingly, [28] generally describe autonomous systems as systems with the ability to process tasks on their own without human influence. Beside the independent task fulfilment, the high adaptability to changing environments is one major characteristic.

In contrast, [29] interpret autonomous systems from an industry perspective in the context of industrial automation systems as the highest level of automation. In this regard, the authors describe an automation system characterized by little to no human influence while system tasks are pre-defined using a predetermined rule-based decision-making in structured environments. Autonomous systems, on the other hand, are described by learning-based capabilities and the ability to adapt to changing system conditions while actions are not pre-programmed. Complementary, [30] describe autonomy in the context of industrial automation systems by four major characteristics commonly used in definitions: First, a systematic process execution is stated which is defined as the ability of a system to execute modeled processes. Second, the adaptability to changing environments for reaching its goals is mentioned. Furthermore, self-governance as the system's ability to manage its resources without human intervention through context-awareness and self-containedness of the system (defined goal and scope of the system) are stated. In the authors perspective, the autonomous system is an extension of the (intelligent) automation system by the above-mentioned further characteristics. Here, self-x capabilities are considered as characteristics of autonomous systems but as-well of automation systems depending on the specific self-x property.

An overview of essential self-x capabilities for cyber-physical systems (CPS) is given by [31]. In this respect, self-x is described as e.g. self-description, self-organization, self-control and self-configuration. All relevant self-x capabilities are ordered within a hierarchy while the authors allocate these capabilities to levels of autonomy. As a result, autonomy is described by these self-x capabilities which enable a certain level of autonomy while an increasing level of autonomy comes along with a decrease in human control. Nevertheless, in line with [30], non-autonomous systems as well can be characterized by certain self-x capabilities as for example self-description. Self-x capabilities are not solely part of autonomous systems but depend on the self-x characteristic and might also describe automation systems with less or no autonomy.

In contrast, the term self-organization on the one hand can be a representative of self-x and on the other hand can be regarded as a separated concept as in [20]. The authors define self-organizing systems as collection of processes of decentral decision-making in heterarchical structures that require the ability for autonomous decisions of interacting entities. In conclusion, the authors see autonomy as a part of the concept of self-organization. [22] acquire a differentiation between the term self-organization on the one hand and autonomous control on the other hand. Here, self-organized systems are regarded as the ability of a system to "design its processes und systematic structures in an autonomous manner" and is therefore more focused to an organizational level. Whereas autonomous control is considered according to [27] and is regarded on an execution level or single object level of the corresponding system. [31] in contrast, consider self-organization as one self-x capability of the highest level of autonomy.

In conclusion, the above-mentioned literature describes autonomy to a certain degree in a similar way, but some inconsistencies and differences can be identified especially in the differentiation with autonomy and automation as well as the terms autonomy and self-x. These are summarized within Figure 1.

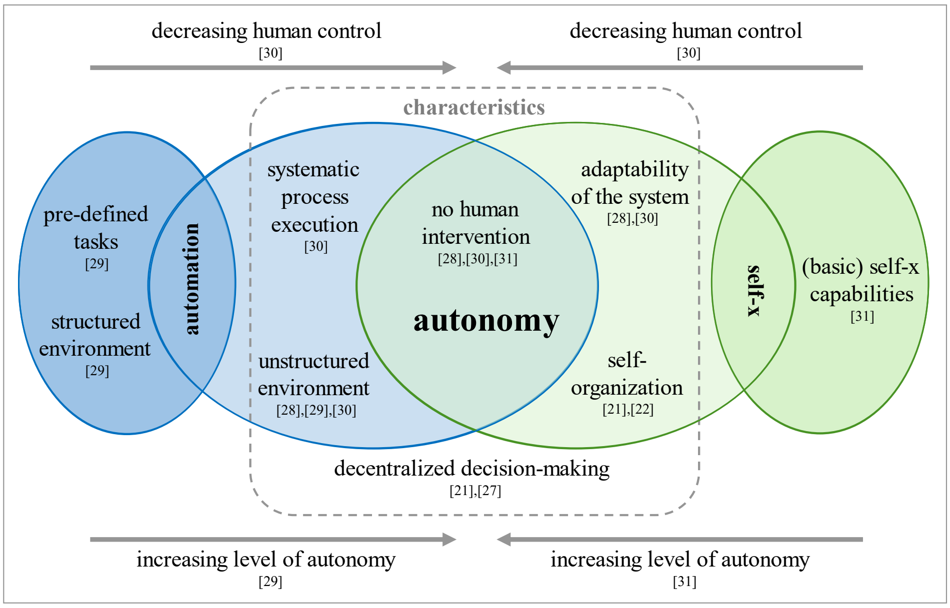


Figure 1: Summary of term differentiation in literature referring to system's characteristics

Furthermore, as most of the sources focus on production systems in general, a specific definition of the term autonomy in context of production logistics needs to be derived dissolving the described inconsistencies (cf. chapter 3.2).

3.2 Definition of autonomy in production logistics

Based on the section above, a definition for autonomy in production logistics can be derived. For a better understanding of the main definition elements, a more detailed explanation will follow below. In this paper, autonomy in production logistics is defined as follows:

> *"Autonomy is the ability of a system to make decentralized decisions* ***without human intervention*** *in order to reach pre-defined goals (transport tasks) and to cope with an* ***uncertain, unknown, and/or dynamically changing environment****. Therefore, the transport task fulfilment related to logistic-specific objectives is realized through an* ***internal intelligence*** *of cooperating autonomous resources."*

Within the mentioned decentralized decision-making, decision-problems are split into smaller problems and only local information depending on the systems environment is considered [20,27]. Therefore, information is generated and processed by the individual system's resources itself. This comes along with the absence of human intervention as the decision-process is realized without external trigger or control by humans [31]. Consequently, the system as well as all individual resources have the ability of self-organization and self-adaption. As part of the decision-making process the achievement of pre-defined system objectives i.e., specific key performance indicators, is pursued. In relation to that, resource tasks like route planning, collision avoidance or navigation in the context of production logistics need to be fulfilled in alignment with the overall system objectives (goal-orientation) while task allocation is again achieved without external control [30]. Especially the cooperation and interaction of individual autonomous resources is required to realize the above explained elements such as decision-making and task fulfilment. In this context, the required adaptability, and the ability to learn for optimized decision-making is realized by internal intelligence of these cooperating autonomous resources.

As in practical not all elements of the definition are fulfilled by every autonomous vehicle, the classification as "autonomous" is insufficient and does not help when comparing AGVs with varying characteristics. That's why different levels of autonomy need to be considered and are described in the next section.

4. Framework for levels of autonomy in production logistics

Based on the above definition of autonomy for production logistics (cf. chapter 3.2) it is possible to specify a description of a universal framework for classifying AGVs regarding their level of autonomy. This framework helps to create comparability and to simplify the choice of an appropriate AGV for a production system by linking skills (cf. chapter 4.1) and tasks (cf. chapter 4.2) in a standardized way. Because in production logistics, there is not necessarily a human worker involved in the task fulfilment the way of cooperation between human and system cannot present a valid classification criterion as it is done for autonomous vehicles in the automotive context (see definitions by National Highway Traffic Safety Administration and Society of Automotive Engineers). The framework presented hereinafter (cf. chapter 4.3) aims to answer the question how to classify the level of autonomy of an autonomous system implemented in a production system.

4.1 Skills of AGVs

The characteristics of AGVs define the skills they can offer to fulfil tasks and influence therefore their level of autonomy. The skills of an AGV depend on the hardware and software components the manufacturer has implemented. In production logistics, as explained above, we consider a technical view "without" implication of any human. An analysis of available publications showed that there exist various approaches to classify skills of autonomous vehicles. The framework is based on the work of [33,32,34] and the following five main skills defining autonomy in the context of production logistics are derived:

- Acquisition of information, i.e. collecting data via various channels/ways
- Information processing, i.e. generating knowledge out of the collected data
- Decision making, i.e. choosing what to do based on the derived knowledge
- Interaction, i.e. communicating with the environment for the execution of a task
- Control, i.e. checking and documenting the successful execution of a task

These five skills constitute the foundation for AGVs being able to fulfil tasks that occur when these AGVs are used in production logistics. The characteristics of AGVs allow to clearly delimit the scope as well as the content of each of these skills and therefore, they have been chosen for the framework presented in this paper. Summing up, the skills of the AGVs define which role an autonomous transport system can take in the production system, i.e. how responsibilities can be shared with an external system (cf. chapter 4.3).

4.2 Tasks of AGVs in production logistics

When analysing the role of logistics in production environments and comparing different approaches (cf. chapter 2.1), four central tasks of autonomous resources can be derived that AGVs have to complete and which are relevant for defining autonomy: navigation, task assignment, collision avoidance, and charging. In order to clarify the scope of each of these tasks as well as their meaning in this paper and hence for the presented framework the four tasks are described below:

- Navigation: This task comprises in particular registering the existing production layout, implementing strategies for how to reach a destination in the production layout using a given algorithm and documenting current routes as well as locations of moving vehicles [25].
- Task assignment: The basis of transporting materials, semi-finished products or finished products consists in deciding which transport resource fulfils which transport task considering defined rules. A production planning system collects all the tasks and disposes of supplementary information like work process, specific requirements, and due dates.
- Collision avoidance: While moving the autonomous vehicle has to consider and avoid collisions with either potentially moving objects, i.e. other vehicles or humans, or static objects, i.e. “things” standing around, that are not captured in the production layout. At crossings there need to be strategies on how to assign priorities in order to avoid dead locks. For more information on the classification of obstacles and the choice of a strategy in the case of collisions like waiting or taking alternative routes see [35].
- Charging: This approach does not focus on strategies for charging (cf. other publications), but on the influence of this procedure in logistics as it interrupts the workflow and is consequently relevant for planning and control. Here, only the supervision of battery charge is taken into account.

The developed framework (cf. chapter 4.3) is based on these task descriptions as it is fundamental for any kind of standardized approach to dispose of a clear definition of the applied basis. Their extent is consciously limited to the jobs that can be assigned to an autonomous transport system applied in production logistics of manufacturing companies.

4.3 Description of the framework

Based on the skills and tasks described in chapter 4.1 and 4.2, the framework proposed hereafter combines these aspects. There can be three ways of distributing the four tasks between the autonomous transport system and an external system (human and/or IT system) based on the five skills:

(1) No external control (except initial order registration), i.e. the autonomous system proposes all necessary skills
(2) Implication of external system, i.e. division of responsibilities and the external system is only responsible for the initial acquisition of information
(3) Control via external system, i.e. the autonomous system proposes only the "executing" skill of interaction

Table 2 specifies the three possible ways for distributing responsibilities between the autonomous transport system and an external system for the five skills (cf. chapter 4.1).

Table 2: Possibilities for distribution of responsibilities

	(1)	**(2)**	**(3)**
Acquisition of information	Autonomous transport system	External system	Autonomous transport system
Information processing		Autonomous transport system	External system
Decision making			
Interaction			Autonomous transport system
Control			External system

The work presented in chapter 4.1 and 4.2 is transferred into a framework by considering these three ways of distributing tasks. In theory, for each of the four tasks an AGV can take over each of the five skills either completely on its own, partly with an external system or transfer it to the external system, i.e. three possible levels per task as introduced above. In practice, not for every task every way of responsibility for the skills is reasonable, so the choices have to be reduced:

- Navigation: (1), (2), (3), i.e. all three ways of responsibility are possible
- Task assignment: (2), (3), i.e. an external system is always required
- Collision avoidance: (1), (3), i.e. the acquisition of information (concerning obstacles) and the interaction is completed by the autonomous transport system itself
- Charging: (1), (3), i.e. the acquisition of information (concerning charging level) is completed by the autonomous transport system itself

As mentioned, not all ways of responsibility are applicable for the four tasks when defining levels of autonomy. This results in three ways for the navigation task and two ways respectively for task assignment, collision avoidance and charging. When additionally considering dependences between the tasks especially between navigation and collision avoidance which are linked to the strategies implemented in AGVs for these tasks, eleven levels of autonomy can be distinguished. They arise from three possible combinations between navigation and collision avoidance, two ways for task assignment and two ways for charging:

$$11\ levels\ of\ autonomy = 3 \cdot 2 \cdot 2 - 1 \tag{1}$$

One possibility has to be subtracted for the combination when all tasks are executed by an external system except charging. This would not be reasonable.

Figure 2 summarizes the approach for the definition of eleven levels of autonomy in production logistics.

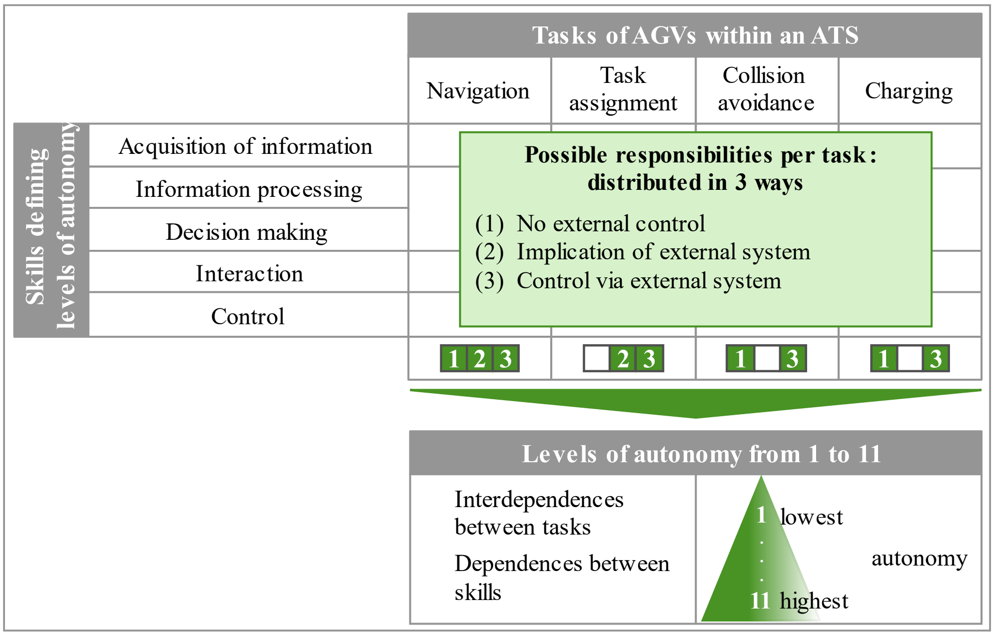


Figure 2: Framework for the classification of autonomy in the context of production logistics

For arranging the levels of autonomy, the rule applies “the less implication of an external system the more autonomy of a system” as proposed also e.g. by [32]. The order was determined by calculating the distance to the origin in a 3D coordinate system by using the number of implicated external systems for the skills acquisition of information, information processing, decision making, interaction, and control. The closer to the origin the more autonomous the autonomous system is. The numbering was inverted compared to the description in [32], so level 11 describes the highest and level 1 the lowest level of autonomy.

5. Conclusion

After shortly introducing the relevant state of the art regarding production logistics and autonomous transport systems, this paper presents an analysis of the current use of the terms autonomy, automation and self-x-approaches in production logistics. These are compared and differentiations in the terminology are summarized. As there exists no clear characterisation of autonomy for production logistics, a definition for this important term in the context of the rising use of AGVs in manufacturing environments is deduced. Afterwards, a framework for classifying AGVs based on their skills and the tasks they have to fulfil is explained. This approach differentiates between eleven levels of autonomy in production logistics.

The presented framework is necessary for decision-makers in manufacturing companies in order to choose in a next step an appropriate AGV for a production system and its specific characteristics. Therefore, the framework provides a basis and is part of a procedure for the organisational integration of an autonomous transport system in a production system. This aspect becomes more and more important due to the rising demand of flexibility in transportation systems and the request for the use of autonomous systems. Further research has to be done on the relation between vehicles classified with the proposed levels of autonomy and their appropriate use in different production organisations.

References

[1] Reinhart, G. (Ed.), 2017. Handbuch Industrie 4.0: Geschäftsmodelle, Prozesse, Technik. Hanser, München, 1774 pp.

[2] Scholz-Reiter, B., Rekersbrink, H., Görges, M., 2010. Dynamic flexible flow shop problems—Scheduling heuristics vs. autonomous control. CIRP Annals 59 (1), 465–468.

[3] Dais, S., 2014. Industrie 4.0 - Anstoß, Vision, Vorgehen, in: Bauernhansl, T., Hompel, M. ten, Vogel-Heuser, B. (Eds.), Industrie 4.0 in Produktion, Automatisierung und Logistik. Anwendung, Technologien, Migration. Springer Vieweg, Wiesbaden, pp. 626–634.

[4] Günthner, W., Klenk, E., Tenerowicz-Wirth, P., 2017. Adaptive Logistiksysteme als Wegbereiter der Industrie 4.0, in: Vogel-Heuser, B., Bauernhansl, T., Hompel, M. ten (Eds.), Handbuch Industrie 4.0. Bd. 4: Allgemeine Grundlagen, 2. Auflage ed. Springer Vieweg, Berlin, pp. 99–125.

[5] Schuhmacher, J., Baumung, W., Hummel, V., 2017. An Intelligent Bin System for Decentrally Controlled Intralogistic Systems in Context of Industrie 4.0. Procedia Manufacturing 9, 135–142.

[6] Fottner, J., Clauer, D., Hormes, F., Freitag, M., Beinke, T., Overmeyer, L., Gottwald, S.N., Elbert, R., Sarnow, T., Schmidt, T., Reith, K.B., Zadek, H., Thomas, F., 2021. Autonomous Systems in Intralogistics – State of the Art and Future Research Challenges. Logistics Research.

[7] Delfmann, W., ten Hompel, M., Kersten, W., Schmidt, T., Stölzle, W., 2018. Logistics as a science: Central research questions in the era of the fourth industrial revolution: Logistics Research Iss. 9, Bremen, 13 pp.

[8] Riechmann, C., 1998. Reorganisation im Logistikbereich zur Verringerung der Durchlaufzeiten und Lagerbestände: Am Beispiel des Logistikzentrums des Automobilzulieferers HÜCO electronic GmbH. Diplomarbeiten Agentur, Hamburg, 1 p.

[9] Grundstein, S., Schukraft, S., Görges, M., Scholz-Reiter, B., 2013. Interlinking central production planning with autonomous production control, in: Marascu-Klein, V. (Ed.), Advances in Production, Automation and Transportation Systems. WSEAS Press, Brasoc, Romania, pp. 326–332.

[10] Dumitrescu, R., Gausemeier, J., Slusallek, P., Cieslik, S., Demme, G., Falkowski, T., Hoffmann, H., Kadner, S., Reinhart, F., Westermann, T., Winter, J., 2018. Studie "Autonome Systeme": Studien zum deutschen Innovationssystem Nr. 13-2018.

[11] Besenfelder, C., Brüggenolte, M., Austerjost, M., Kämmerling, N., Pöting, M., Schwede, C., Schellert, M., 2017. Paradigmenwechsel der Planung und Steuerung von Wertschöpfungsnetzen.

[12] Pfohl, H.-C., 2018. Logistiksysteme: Betriebswirtschaftliche Grundlagen, 9. Aufl. 2018 ed. Springer Vieweg, Berlin, Heidelberg, 437 pp.

[13] Lödding, H., 2019. Produktionslogistik, in: Furmans, K., Kilger, C. (Eds.), Betrieb von Logistiksystemen, vol. 35. Springer Berlin / Heidelberg, Berlin, Heidelberg, pp. 107–131.

[14] Schuh, G., Brandenburg, U., Liu, Y., 2015. Evaluation of Demand Response Actions in Production Logistics. Procedia CIRP 29, 173–178.

[15] VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA), 2013. Cyber-Physical Systems: Chancen und Nutzen aus Sicht der Automation. Thesen und Handlungsfelder.

[16] Schuh, G., Stich, V., 2012. Produktionsplanung und -steuerung 1: Grundlagen der PPS, 4. Aufl. ed. Springer, Berlin.

[17] Wiendahl, H.-P., 1997. Fertigungsregelung: Logistische Beherrschung von Fertigungsabläufen auf Basis des Trichtermodells. Hanser, München, 382 pp.

[18] Lödding, H., 2016. Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration, 3. Aufl. ed. Springer Vieweg, Berlin, Heidelberg.

[19] Grundstein, S., Schukraft, S., Scholz-Reiter, B., Freitag, M., 2015. Evaluation System for Autonomous Control Methods in Coupled Planning and Control Systems. Procedia CIRP 33, 121–126.

[20] Scholz-Reiter, B., Höhns, H., 2006. Selbststeuerung logistischer Prozesse mit Agentensystemen, in: Schuh, G. (Ed.), Produktionsplanung und -steuerung. Grundlagen, Gestaltung Und Konzepte. Springer, Dordrecht, pp. 745–780.

[21] Schwarz, C., Schachmanow, J., Sauer, J., Overmeyer, L., Ullmann, G., 2013. Selbstgesteuerte Fahrerlose Transportsysteme, 8 pp.

[22] Schuhmacher, J., Hummel, V., 2020. Self-organization and autonomous control of intralogistics systems in line with versatile production at Werk 150.

[23] Scholz, M., 2019. Intralogistics Execution System mit integrierten autonomen, servicebasierten Transportentitäten.

[24] Nyhuis, P., Wiendahl, H.-P., 2012. Logistische Kennlinien: Grundlagen, Werkzeuge und Anwendungen, 3. Aufl. 2012 ed. Springer, Berlin, Heidelberg.

[25] Ullrich, G., 2014. Fahrerlose Transportsysteme: Eine Fibel - mit Praxisanwendungen - zur Technik - für die Planung, 2., erw. und überarb. Aufl. ed. Springer Vieweg, Wiesbaden, 241 pp.

[26] VDI e.V. VDI-Richtlinie, 2510: Fahrerlose Transportsysteme (FTS).

[27] Windt, K., Böse, F., Philipp, T., 2008. Autonomy in production logistics: Identification, characterisation and application. Robotics and Computer-Integrated Manufacturing 24 (4), 572–578.

[28] Dumitrescu, R., Westermann, T., Falkowski, T., 2018. Autonome Systeme in der Produktion. I40M 2018 (6), 17–20.

[29] Gamer, T., Kloepper, B., Hoernicke, M., 2019. The way toward autonomy in industry - taxonomy, process framework, enablers, and implications, in: IECON 2019 - 45th Annual Conference of the IEEE Industrial

Electronics Society. IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, Lisbon, Portugal. IEEE, pp. 565–570.

[30]Müller, M., Müller, T., Ashtari Talkhestani, B., Marks, P., Jazdi, N., Weyrich, M., 2021. Industrial autonomous systems: a survey on definitions, characteristics and abilities. at - Automatisierungstechnik 69 (1), 3–13.

[31]Stock, D., Bauernhansl, T., Weyrich, M., Feurer, M., Wutzke, R., 2020. System Architectures for Cyber-Physical Production Systems enabling Self-X and Autonomy, in: 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria. IEEE, pp. 148–155.

[32]Castelfranchi, C., Falcone, R., 2003. From Automaticity to Autonomy: The Frontier of Artificial Agents, in: Weiss, G., Hexmoor, H., Castelfranchi, C., Falcone, R. (Eds.), Agent Autonomy, vol. 7. Springer US, Boston, MA, pp. 103–136.

[33]Beer, J.M., Fisk, A.D., Rogers, W.A., 2014. Toward a framework for levels of robot autonomy in human-robot interaction. Journal of human-robot interaction 3 (2), 74–99.

[34]Parasuraman, R., Sheridan, T.B., Wickens, C.D., 2000. A model for types and levels of human interaction with automation. IEEE transactions on systems, man, and cybernetics. Part A, Systems and humans : a publication of the IEEE Systems, Man, and Cybernetics Society 30 (3), 286–297.

[35]Krä, M., Vogt, L., Spannagl, V., Schilp, J., 2020. Multi-agent path planning: comparison of different behaviors in the case of collisions, in: Schüppstuhl, T., Tracht, K., Henrich, D. (Eds.), Annals of Scientific Society for Assembly, Handling and Industrial Robotics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 217–227.

Biography

Michaela Krä (*1990) has been working at the Chair of Digital Manufacturing, Faculty of Applied Computer Science at Augsburg University since 2016. In 2015, she completed her bilingual studies in “Production and Automation” in Munich and Paris. She acquired practical knowledge among others in internships at MINI (BMW Group, Great Britain), MTU Aero Engines (Germany) and Airbus (France).

Leonard Eckart (*1997) has been working at the Chair of Digital Manufacturing, Faculty of Applied Computer Science at Augsburg University since 2021. During his studies in Industrial Engineering at Augsburg University starting in 2015 he gained practical experience among others in the automotive industry at BMW Group, Audi Brussels (Belgium) and Faurecia Clean Mobility.

Marinella Rito (*1993) has been working as a ME Engineer Automation & Design for Webasto Roof & Components SE since February 2021. From 2014 to 2021 she studied Industrial Engineering with focus on "Management and Sustainability" at Augsburg University. During her studies, she gained practical experience as an intern at MAN Truck & Bus AG and Webasto Roof & Component SE.

Johannes Schilp (*1976) has been head of the Chair of Digital Manufacturing, Faculty of Applied Computer Sciences at Augsburg University since 2015. Prof. Dr.-Ing. Johannes Schilp is also head of the processing department at the Fraunhofer Research Institution for Casting, Composite and Processing Technology and member of scientific networks like WG MHI and automation research groups (VDMA).

3rd Conference on Production Systems and Logistics

Towards A Digital Workflow Solution For Cradle-To-Gate Sustainability Information In Textile Value Chains

Dieter Stellmach[1], Michael Weiß[1], Jürgen Seibold[1], Meike Tilebein[2, 1]
[1]*German Institutes of Textile and Fiber Research, Denkendorf, Germany*
[2]*University of Stuttgart, Institute for Diversity Studies in Engineering, Stuttgart, Germany*

Abstract

Sustainability aspects and their verification are becoming indispensable for companies in the textile industry from both an economic and a legal perspective. The reason for this is that there is a large number of different certificates, specifications, and labels, such as Global Organic Textile Standard, Fairtrade, or OekoTex, as well as legislation, such as the German Act on Corporate Due Diligence Obligations in Supply Chains issued in 2021.

Hence, the requirements for keeping the proof, e.g. for the batch-accurate world-wide tracing of organic cotton for clothing, or for the necessary transparency to determine the carbon footprint or the recycling percentage, are becoming more and more associated with considerable effort, especially for small and medium-sized enterprises (SMEs). Depending on the certificate or specification, SMEs need not only to determine their own sustainability information (gate-to-gate), but also that of the upstream stages of the value chain (cradle-to-gate). The multi-stage value chains of the SME-dominated textile industry, together with the vast and fast-changing variety of materials and products, lead to high complexity in processes and communication. In addition, when confronted with batch-related sustainability criteria and a variety of sustainability and labelling requests from different customers, SMEs have to spend an increasing amount of time and effort on the reliable provision and communication of the respective information.

The paper describes the challenges and existing approaches, e.g. the use of blockchain technology, associated with the provision of cradle-to-gate sustainability information in textile SMEs and proposes a holistic framework enabling SMEs along the value chain to configure and implement an infrastructure for efficient, fully digital cloud-ready workflow, based on process models and textile product master trees, in order to address these challenges.

Keywords

Cradle-to-Gate Sustainability; Textile Industry; SME; Blockchain

1. Challenges of sustainability information in textile value chains

Sustainability in social, economic and ecological terms is one of the megatrends of recent years [1]. In the textile industry sustainability is also becoming increasingly important for purchasing decisions [2] in the B2B and B2C sectors, as shown in the following examples:

- Large quantities of textiles end up in residual waste and thus pollute the environment. Closed material cycles and recycling have not been established to a larger extent so far. One reason for this is the lack of information on which materials have been processed in the product and how they can

DOI: https://doi.org/10.15488/12155

be separated, reprocessed and recycled. The provision of such information along the entire value chain is possible [3, 4], but still a challenge.

- The working conditions of the textile industry in low wage countries often do not meet the standards demanded by consumers and NGOs. In addition to legislation, there are standards, certificates and labels to prove social sustainability, but these require a reliable and transparent data basis across the entire value chain. Providing this data is time-consuming, especially when it comes to specific batches and products.

Today, (end) customers expect or demand access to comprehensive data, e.g. via an identifier (QR code, RFID or similar), on environmental impacts (such as carbon footprint), on social standards during production or on the material composition of textile products including all preliminary products [5, 6]. Expectations vary widely across customers, which is reflected in the large number of sustainability requirements and challenges that have to be met by a single company within the textile value chain [7]. Additionally, legislation puts more and more focus on sustainability aspects, e.g. in Germany [8].

Sustainability cannot be measured with a single indicator. Rather, there is a multitude of criteria of different types and structures and thus many types of certificates, labels and standards with different requirements, criteria and system boundaries. SMEs in the textile industry therefore have to provide their data in ever new forms. For example, the producer and all suppliers need a Scope certificate [9] for each delivery of a certified product according to the Global Organic Textile Standard (GOTS) [10]. Transaction certificates for the corresponding batches must also be submitted for the product and all pre-stage products. This process is time-consuming, cost-intensive and limited to specific statements. This form of cross-company communication can only be implemented for a limited number of sustainability criteria by an SME without a supporting organisational and information technology infrastructure.

Traceability for sustainability purposes is the focus of such an infrastructure. Traceability includes the objectives of transparency, identification, authenticity and quantification [11], and corresponding information on: (a) origin of the relevant materials back to the source, (b) material composition (especially with regard to bio-materials and recycling), (c) ecological life cycle assessment parameters (carbon footprint, blue water consumption) with differentiation according to Scope 1 (direct emissions), Scope 2 (indirect emissions through energy used, especially electricity) and Scope 3. 1 (emissions from input products, services) [12] and (d) ecological and ethical verification (certificates such as GOTS [9] or IVN BEST [13]).

In the context of the preparation of internal company data [14], textile companies are faced with some challenges. The type of information collected from suppliers must be selected and for communication the form, granularity and participants must be determined. Also, practical issues like aggregation and allocation of sustainability data on production processes and products arise. Finally, the textile companies have to protect and secure their data. There are numerous individual systems, e.g. Blockchain [15], for this, but no holistic solution. The reason for this is that textiles comprise a variety of products or product types, from clothing to home textiles to technical textiles, with different requirements for sustainability information. Production is usually multi-stage and distributed across many different companies along the value chains with a wide range of semi-finished and raw materials being used. This results in complex material flow with a diversity of textile traceability targets and certificates. The highest complexity in the requirements arises in the case of cross-stage traceability at batch level. Here, the material interrelationships and the associated sustainability information must be mapped across the entire value chain. For one product, the product family tree (see Figure 1) illustrates the basic problem with traceability. Materials from numerous preliminary stages in different quantities and of different natures go into a finished product, such as a garment. For complete traceability, this situation must be mapped in terms of information technology and supported across companies.

However, a reliable provision of information across all levels with the required granularity and the dynamic of the textile value chains with changing network structures from batch to batch (see example in Figure 1 with different yarn networks) is time-consuming. Corresponding data preparation and allocation are today often performed manually and using individually created, complicated and therefore error-prone spreadsheet models. As a result, the quality and integrity of the information also suffer.

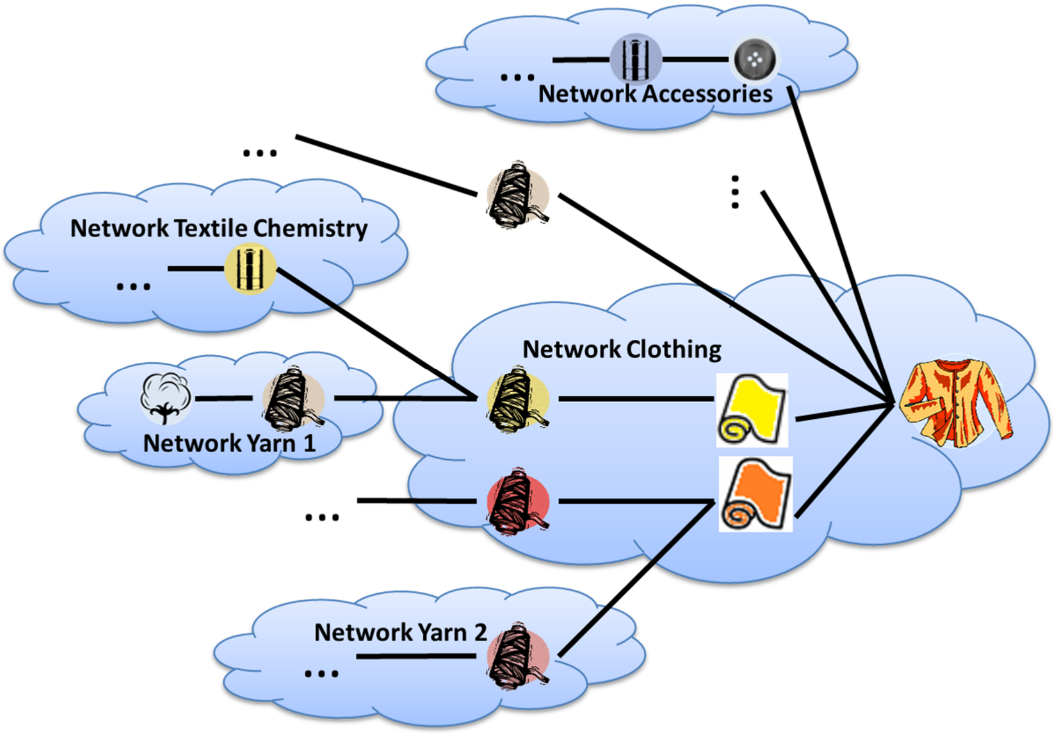


Figure 1: Example of a textile product master tree

2. Solutions for sustainability information and open issues

To cope with the complexity of tracking and tracing of products and materials as well as communicating sustainability information along the value chain a large variety of solutions have been proposed. These solutions often include one or more of the following aspects:

- Closed value networks
- Cross-level communication of sustainability information
- Provision of sustainability information
- Conformity of sustainability information

However, the specific demand of textile value chains is often only partially covered by these solutions like explained in the following sections.

2.1 Closed value networks

There are solutions for closed value networks in which all partners work exclusively according to a common procedure or even standard. ‘bioRe® Sustainable Textiles’ [16], for example, allows each product to be traced back up to the cotton field via a code in the garment. Authenticity is guaranteed by completely organising and controlling the value chain and making the information available to consumers on proprietary platforms. These approaches are also suitable for multi- or fully integrated companies that can manage sustainability information internally within their own information architecture. Furthermore, there are software solutions that enable traceability. However, these often only cover individual aspects of the above requirements. One example is the software of the start-up ‘sustainabill’ with a cloud platform for uncovering and analysing transparency along supply chains [17].

Since most companies in the textile industry do not operate in closed value networks and available software systems only cover partial aspects of traceability (e.g. transparency and authenticity at [16]), there is a gap here. There is no systematic and flexible support for the cross-stage provision and communication of sustainability information for the textile industry. Companies cannot meet the obligations and expectations due to a lack of knowledge about their supply chains and poor information quality [18].

2.2 Cross-level communication of sustainability information

There are several quasi-standards from the GS1 organisation for communication. The ‘Global GS1 Traceability Standard’ is a very comprehensive framework document for the design of interoperable traceability systems in supply chains [19]. Based on globally unique product numbers GTIN and location numbers GLN [20], end products can first be identified and tracked, typically between end manufacturer and retailer. Pre-products back to the first raw material source are also considered. Building on this, Global Textile Scheme has developed [21]. This initiative aims to simplify and standardise the exchange of data, especially sustainability data, in the textile value chain.

For the communication of master data, there is the Global Data Synchronisation Network (GDSN). However, this does not allow traceability at batch level. A comprehensive implementation guide complements the formal GS1 standards of the GDSN with advice on their implementation and operation [22]. With the GDSN, product content is automatically uploaded to data pools, maintained and shared. Only implicit communication is used, i.e. the partners independently retrieve the data they need from the master data provided. The largest data pool provider for this is the company Atrify [23] in Germany. For the textile flooring sector, there is the European Product Information System [24]. For the communication of batch-based information for the traceability of textile products, there are some basics but no concept analogous to master data or as in other sectors, such as the food industry [25].

2.3 Provision of sustainability information

Today, the internal processing of data for traceability at batch level is usually carried out in textile-specific ERP solutions. However, these cannot provide the functionalities required for the cross-stage provision of sustainability information, such as (1) uniform identifiers for batches, (2) their unambiguous mutual assignability with (3) a high degree of security and integrity. Within the company, individual identifiers are usually used for this purpose, which are also communicated to customers. For external communication, there are, for example, independent formats such as the "Universally Unique Identifier" from the Open Software Foundation as part of the "Distributed Computing Environment" [26] in addition to the proprietary GS1 quasi-standard GTIN for the unique identification of objects.

For companies, the security and integrity of their externally communicated information is particularly important. One promising approach to this is blockchain technology [27]. In addition to the first blockchain approaches for cryptocurrencies [28, 29], there are currently open-source alternatives for private company networks (Quorum [30], Corda [31], Hyperledger Fabric [32]). The storage of sustainability data was also investigated [33, 34]. In several scientific publications, the potentials of blockchain in logistics, the food industry and healthcare were presented [25, 35]. Furthermore, Amazon Web Services offers solutions based on Hyperledger Fabric for the textile industry, among others [36]. There are also efforts to store information of the global standard GS1 in blockchains [37, 38]. The potential of blockchain for traceability for the textile industry has already been recognised [27, 39]; numerous companies can show pilot projects in this regard (IBM [40], Provenance [41], Lenzing [42]), also by means of the blockchain alternative IOTA [43] for textiles [44].

2.4 Conformity of sustainability information

Conformity with customer requirements or certificates is another aspect that poses major challenges for SMEs in the textile industry. For example, if a company wants to be part of a "CO2-neutral supply chain", you have to manage your Scope3.1 CO2 emissions [12, 45]. The necessary Scope2 information from the own company is only correctly available for a few companies today, whereas the necessary Scope3.1 data of the purchased raw materials [12,45] are, as of today, not available in most cases. The conclusion of the practical recommendations for data collection and calculation of greenhouse gas emissions in the supply chain clearly shows this: 'Even among the German companies on the CDP A-List, hardly any are able to determine their Scope3.1 emissions on the basis of primary data. Estimates based on average industry data and EEIO databases or tools predominate' [46]. A correct batch-related reporting of the CO2 quantity on Scope3.1 level, supplemented by various sustainability information of certificates by means of transaction certificates, is not possible for textile companies with the communication content and structures available today.

3. A framework for sustainability information in the textile industry

As shown above, there is a wide range of solutions for traceability and communication of sustainability information. However, they don't fully match the requirements of companies of the textile industry, especially of SMEs. The restriction to certain types of sustainability information, e.g. only support of specific labels, the lack of support for open value networks, or limited support of tracking and tracing on batch level limit the suitability of the solutions.

For a systematic and flexible support of SMEs in the textile industry, the authors propose a holistic framework for cradle-to-gate sustainability information. The framework consists of three layers (see Figure 2) and is based on a cloud approach. The top layer supports the breakdown of specific sustainability enquiries into basic enquiry types. This makes it very easy to integrate new types of enquiries and does not limit the framework to specific certificates or labels.

The next layer provides the methodological support to answer the basic enquiries. The textile industry domain workflow, based on the product master tree of textiles, implements the necessary query steps that are required for the respective basic enquiry types. Thus, the framework supports open value networks. This is supplemented by transformation rules that provide the necessary functions, e.g. conversion of units or breakdown of data to batch level, for the preparation of the raw data. Conformity rules ensure that all requirements of certificates and labels are met. Intermediate and final enquiry results but no raw data are stored using a generic, type specific information structure. This also fulfils the requirement for non-disclosure of company confidential data.

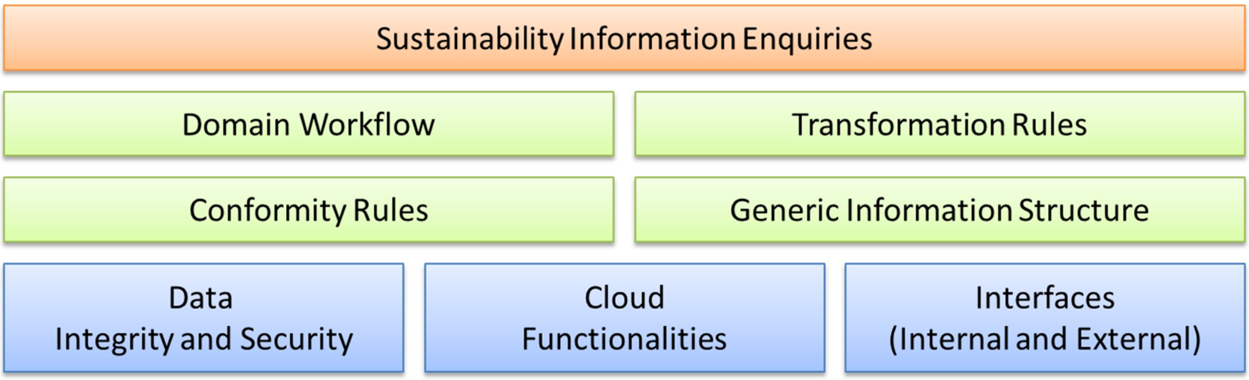


Figure 2: Architecture of the sustainability information framework for the textile industry

The bottom layer provides the technological support required for answering sustainability enquiries. Data integrity and security technologies ensure the safety and validity of company data. Cloud functionalities enable the required communication and transformation activities. Finally interfaces to cloud-internal companies and to other sustainability information networks complement the framework.

Sustainability Information Enquiries: In order to support a wide range of labels and certificates, each specific enquiry must be traced back to basic types. For this mapping, a classification system is necessary that describes the requirements of each enquiry and thus clearly assigns them to a basic type. One example is the system boundary of the enquiry with possible characteristics such as gate-to-gate and cradle-to-gate. However, the classification is not only used to identify a basic enquiry type, it is also used to configure it, i.e. to adapt the workflow, the transformation rules and the conformity rules to the specific request belonging to the basic enquiry type.

Domain Workflow: The textile domain-specific workflows describe the necessary process steps for each basic enquiry type. These can still be adapted to the certificates and labels based on the classification of the previous layer. The basis for the final workflow is the product master data of the product for which the enquiry requests sustainability information. The product master data together with the batch numbers clearly identify the network to which the request refers. This means that not only static networks are supported but also dynamic ones, which are the norm in the textile industry.

Transformation Rules: In addition to the pure collection of data, which is organised by the domain workflow, the raw data must also be processed. The way in which raw data is processed is determined by the transformation rules. As with the domain workflow, the classification from the top level determines the required transformations and can also be used to configure individual transformations. For example, the transformation rules determine how values are to be recorded (industrial mean, annual mean, mean value in the production period, exact values).

Conformity Rules: Finally, the conformity rules comprise framework conditions of certificates and labels that cannot be mapped to the domain workflow and the transformation rules. These rules ensure that requirements such as the existence of transaction certificates in GOTS are fulfilled. Thus, they implement special features of certificates and labels that are not covered by the standardised procedure of the basic enquiry types.

Generic Information Structure: In order to answer sustainability requests a variety of sustainability information must be captured and processed in the production network and the result has to be provided to the requester. Therefore, corresponding generic information structures are required for the different basic enquiry types. These then serve to store the sustainability information captured by the workflow and processed by the transformation rules. Information required by conformity rules must also be stored there. No raw data of the participating companies is stored in the generic information structure.

Data Integrity and Security: The focus is on securing and maintaining the integrity of the data when communicating sustainability information across companies and networks. This protection is of central importance for SMEs of the textile industry, since on the one hand access to raw data would allow conclusions to be drawn about production processes and thus affects the company's business secrets, and on the other hand the integrity of the communicated data must be ensured. Established concepts such as encryption or blockchain are used to secure the data.

Cloud Functionalities: In addition to the standard functions of a cloud, further functionalities must be provided. This refers in particular to the necessary functions that realise the concepts and methods of the higher layers of the framework. These functions are conceptually combined in topic-specific libraries.

Interfaces (Internal and External): For identification of sustainability information, concepts and methods for communication across organisational boundaries are needed. Concepts and methods, currently discussed in international standardisation committees, for both technical (e.g. interfaces, protocols) and organisational (e.g. responsibilities, request and response process) aspects of cross-organisational communication will be used. The textile value chain is sometimes very extensive. Therefore, it cannot be assumed that all actors work together in a common cloud. In this case, different data sources can only be integrated into the workflow to a limited extent. To ensure that the entire value chain can still be covered, established concepts for the integration of sustainability information from external data sources (e.g. companies that are not members of the cloud but are part of the production network) are used.

4. Summary and outlook

Summarising, the proposed framework can enable organisations of the textile industry, in particular SMEs, to create an infrastructure for flexible and efficient generation and communication of sustainability information of textile materials and products on batch-level along the value chain.

The proposed framework can serve all relevant aspects of sustainability information for the vast variety of textile materials, products, processes, as well the huge number of different labels and sustainability certificates. It enables companies of the textile industry to react quickly to new sustainability information demands by a reliable provision and communication of the respective information, and thus reduces related time and effort. This is in particular achieved by the holistic structure of the framework and the decomposition of the problem area into eight well defined framework elements.

For further refinement, focus will be put on elaborating and refining the framework concepts and methods and creating an applicable reference framework solution. This includes for example the further specifications of the textile industry domain towards the domain workflow, the identification and generalisation to basic enquiries types, or initial transformation rules. This will be part of upcoming activities in research projects at the German Institutes of Textile and Fiber Research.

Transparency is necessary to improve sustainability. In particular for the transformation of today's linear textile value chains into new circular economy structures [47, 48], transparency is one of the key enablers. If transparency is not available, the existing and emerging recycling technologies [49] or re-purposing of textiles is impossible without knowing what is in the textile and from which origin, for example. Thus, the proposed framework can significantly support the textile industry in their transformation process towards a circular economy and increased sustainability.

References

[1] Gaudig, A., Ebersberger, B., Kuckertz, A. 2021. Sustainability-Oriented Macro Trends and Innovation Types - Exploring Different Organization Types Tackling the Global Sustainability Megatrend. Sustainability 13(21), 11583.

[2] Desore, A., Narula, S., 2018. An overview on corporate response towards sustainability issues in textile industry. Environment, Development and Sustainability, Vol. 20(4)4, Springer Science+Business Media, Dordrecht, pp. 1439-1459.

[3] https://circular.fashion/

[4] https://www.waste2wear.com/

[5] Karthik, T., Gopalakrishnan, D., 2014. Environmental analysis of textile value chain: an overview. Roadmap to sustainable textiles and clothing, pp. 153-188.

[6] United Nations, 2017. Textile4SDG12: Transparency in Textile Value Chains in Relation to the Environmental, Social and Human Health Impacts of Parts, Components and Production Processes.

[7] ELLEN MAC ARTHUR FOUNDATION, 2017. A New Textiles Economy: Redesigning fashions future. https://www.ellenmacarthurfoundation.org/publications/a-new-textiles-economy-redesigning-fashions-future

[8] Bundestag of the Federal Republic of Germany, 2021. Act on Corporate Due Diligence Obligations for the Prevention of Human Rights Violations in Supply Chains. https://www.bmas.de/SharedDocs/Downloads/DE/Internationales/act-corporate-due-diligence-obligations-supply-chains.pdf

[9] Global Standard gGmbH (ed.), 2021. GLOBAL ORGANIC TEXTILE STANDARD (GOTS) VERSION 6.0. https://global-standard.org/images/resource-library/documents/standard-and-manual/gots_version_6_0_en1.pdf

[10] https://global-standard.org/

[11] World Resources Institute, World Business Council for Sustainable Development (ed.), 2015. The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard. https://ghgprotocol.org/sites/default/files/standards/ghg-protocol-revised.pdf.

[12] World Resources Institute, World Business Council for Sustainable Development (ed.), 2011. Corporate Value Chain (Scope 3) Accounting and Reporting Standard. Supplement to the GHG Protocol Corporate Accounting and Reporting Standard. https://ghgprotocol.org/sites/default/files/standards/Corporate-Value-Chain-Accounting-Reporing-Standard_041613_2.pdf

[13] Internationaler Verband der Naturtextilwirtschaft e.V. (ed.), 2018. NATURTEXIL IVN zertifiziert BEST (IVN BEST). Version 6.1. https://naturtextil.de/wp-content/uploads/2018/10/IVN_BEST_Version_6-1_2018.pdf

[14] Fuchs, M., 2019. Die Zusammenarbeit in der globalen Supply Chain der Bekleidungswirtschaft - Eine Fallstudie: Organisation, Management, Näharbeitsplätze. http://medien.bwv-verlag.de/9783830541585_p.pdf

[15] Rosado da Cruz, A., Cruz, E., 2020. Blockchain-based Traceability Platforms as a Tool for Sustainability. ICEIS (2), pp 330-337.

[16] bioRe® Foundation (ed.), 2019. bioRe® SustainableTextilesStandard. Version 0.0 17.06.2019. https://www.biore.ch/wp-content/uploads/bioRe-Sustainable-Textiles-Standard-2019_version-0.0_public-1.pdf.

[17] sustainabill GmbH, 2021. sustainabill: a cloud platform to actively manage the sustainability of your entire supply chain. https://sustainabill.de/

[18] Deloitte (ed.), 2018. Leadership: Driving innovation and delivering impact. The Deloitte Global Chief Procurement Officer Survey 2018. https://www2.deloitte.com/content/dam/Deloitte/at/Documents/strategy-operations/deloitte-global-cpo-survey-2018.pdf

[19] GS1 (ed.), 2017. GS1 Global Traceability Standard GS1's: framework for the design of interoperable traceability systems for supply chains. https://www.gs1.org/sites/default/files/docs/traceability/GS1_Global_Traceability_Standard_i2.pdf

[20] GS1 Germany (ed.), 2015. GS1 Web Vocabulary Standard. Release 1.6.1. https://www.gs1.org/docs/gs1-smartsearch/GS1_Vocabulary_Standard.pdf

[21] Global Textile Scheme Initiative, 2020. https://www.globaltextilescheme.org/

[22] GS1 AISBL, 2020. GDSN 3.1 Trade Item Implementation Guide. Supplements the formal GS1 Global Data Synchronisation Network (GDSN) standards with advice on their implementation and operation, Release 33. https://www.gs1.org/docs/gdsn/tiig/3_1/GDSN_Trade_Item_Implementation_Guide.pdf

[23] atrify GmbH, 2021. Mit GDSN Daten sicher austauschen, synchronisieren und profitieren. https://www.atrify.com/solutions/gdsn-global-data-synchronization-network/

[24] Gemeinschaft umweltfreundlicher Teppichboden e.V., 2020. PRODIS: PRODuct Information System for textile floor coverings. https://gut-prodis.eu/en/info-gut/prodis

[25] Dujak, D., Sajter, D., 2019. Blockchain Applications in Supply Chain. In: Kawa A., Maryniak A. (eds.). Smart Supply Network. Springer International Publishing, Cham, pp. 21–46.

[26] ISO/IEC 9834-8:2014, 2014. Information technology — Procedures for the operation of object identifier registration authorities — Part 8: Generation of universally unique identifiers (UUIDs) and their use in object identifiers. https://www.iso.org/standard/62795.html

[27] El Messiry, E., El Messiry, A., 2018. Blockchain Framework for Textile Supply Chain Management. In: Chen S., Wang H., Zhang L.-J. (eds.). International Conference on Blockchain (ICBC 2018). Springer International Publishing, Cham. pp. 213–227.

[28] Nakamoto, S., 2009. Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf

[29] Wood, D. G., 2019. Ethereum: A Secure Decentralised Generalised Transaction Ledger. github.com/ethereum/yellowpaper/

[30] JP Morgan Chase (ed.), 2016. Quorum Whitepaper. https://github.com/jpmorganchase/quorum-docs/blob/master/Quorum%20Whitepaper%20v0.1.pdf

[31] Brown, R.G. et al., 2016. Corda: An Introduction. https://docs.corda.net/en/pdf/corda-introductory-whitepaper.pdf

[32] Androulaki, E. et al., 2018. Hyperledger Fabric: A Distributed Operating System for Permissioned Blockchains. In: Proceedings of the Thirteenth EuroSys Conference, pp. 1–15.

[33] Kouhizadeh, M., Sarkis, J., 2018. Blockchain Practices, Potentials, and Perspectives in Greening Supply Chains. In: Sustainability, vol. 10, No. 10, p. 3652.

[34] Westerkamp, M., Victor, F., Küpper, A., 2018. Blockchain-based Supply Chain Traceability: Token Recipes model Manufacturing Processes. http://arxiv.org/abs/1810.09843

[35] Bocek, T. et al., 2017. Blockchains everywhere - a use-case of blockchains in the pharma supply-chain. In: 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM), Lisbon, Portugal, pp. 772–777.

[36] Amazon Web Services, 2020. Blockchain for Supply Chain: Track and Trace. https://aws.amazon.com/de/blockchain/blockchain-for-supply-chain--track-and-trace/

[37] The Linux Foundation, 2021. Hyperledger Grid. https://www.hyperledger.org/projects/grid

[38] TextileGenesis, 2021. Textile Genesis. https://textilegenesis.com/

[39] Fibre2Fashion Pvt. Ltd., 2021. Blockchain is Reshaping the World of Textile & Apparel. https://www.fibre2fashion.com/industry-article/8428/blockchain-is-reshaping-the-world-of-textile-apparel

[40] Coinidol, 2021. Applying Blockchain to Fashion & Textile Industry, Made in Italy Project Resumes. https://coinidol.com/blockchain-italy-project/

[41] Project Provenance Ltd., 2021. Increasing transparency in fashion with blockchain. https://www.provenance.org/case-studies/martine-jarlgaard

[42] Ahmed, W., MacCarthy, B., 2021. Blockchain-Enabled Supply Chain Traceability in the Textile and Apparel Supply Chain: A Case Study of the Fiber Producer Lenzing. Sustainability; Vol. 13(19):10496.

[43] Popov, S., 2015. The tangle. https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf

[44] Jegelka, S., 2019. Implementing an IOTA based Supply Chain Documentation. https://medium.com/topocare-x-iota/implementing-an-iota-based-supply-chain-documentation-cd8103bcec46.

[45] World Resources Institute, World Business Council for Sustainable Development, (ed.), 2013. Technical Guidance for Calculating Scope 3 Emissions. Supplement to the Corporate Value Chain (Scope 3) Accounting & Reporting Standard. https://www.ghgprotocol.org/sites/default/files/ghgp/standards/Scope3_Calculation_Guidance_0.pdf

[46] Global Compact Netzwerk Deutschland (ed.), 2019. SCOPE 3.1. Praxisempfehlungen zur Datenerhebung und Berechnung von Treibhausgasemissionen in der Lieferkette. https://www.sustainable.de/wp-content/uploads/2019/12/DGCN-Diskussionspapier-Scope-3-Lieferkette-CO2-Bilanz_2019.pdf

[47] Köhler, A., Watson, D., Trzepacz, S., Löw, C., Liu, R., Danneck, J., Konstantas, A., Donatello, S. & Faraca, G., 2021. Circular Economy Perspectives in the EU Textile sector, EUR 30734 EN, Publications Office of the European Union, Luxembourg.

[48] acatech - Deutsche Akademie der Technikwissenschaften, 2021. Circular Economy Roadmap for Germany. https://www.acatech.de/publikation/circular-economy-roadmap-fuer-deutschland/download-pdf?lang=en

[49] European Commission, Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs, Duhoux, T., Maes, E., Hirschnitz-Garbers, M., et al., 2021. Study on the technical, regulatory, economic and environmental effectiveness of textile fibres recycling: final report, Publications Office.

Biography

Dieter Stellmach (*1958) has been working at the Center of Management Research of the German Institutes of Textile and Fiber Research in Denkendorf (DITF) since 1986 as senior project manager and EU contact point co-responsible for research in organisation and management of sustainable textile value creation and in textile digitisation. He is active member of networks about fibres and textiles on regional, national and European level.

Michael Weiß (*1973) has been working at the Center of Management Research of the German Institutes of Textile and Fiber Research in Denkendorf (DITF) since 2001. Dr.-Ing. Michael Weiß completed his doctorate in 2019 on model-based design of knowledge-intensive value-adding networks in the textile industry. His research includes modelling of textile value chains and information processing.

Jürgen Seibold (*1965) has been working at the Center of Management Research of the German Institutes of Textile and Fiber Research in Denkendorf (DITF) since 1990. He received his Dr.-Ing. on modelling of textile expertise in expert systems in 1995. He has coordinated many regional, national and European research projects. His research includes the development of methods and models for evaluating the material and energy efficiency of textile production processes as a basis for sustainable production.

Meike Tilebein (*1966) has been the Director of the Institute for Diversity Studies in Engineering at the University of Stuttgart since its establishment in 2009. Prof. Dr. rer. pol. Dipl.-Ing. Meike Tilebein also heads the Center of Management Research of the German Institutes of Textile and Fiber Research Denkendorf (DITF) that supports the textile industry’s digital transformation.

3rd Conference on Production Systems and Logistics

Blockchain technology as the backbone of the internet of things – A taxonomy of blockchain devices

Tan Gürpinar[1], Max Austerjost[2], Josef Kamphues[2],
Jonas Maaßen[1], Furkan Yildirim[1], Michael Henke[1]
[1]*Chair of Enterprise Logistics, TU Dortmund University, Dortmund, Germany*
[2]*Fraunhofer Institute for Material Flow and Logistics, Dortmund, Germany*

Abstract

While the number of blockchain ecosystems is growing, enterprises are confronted with the decision on how data can be securely and reliably transferred to blockchains. Even though current blockchain solutions proof to be a secure way for cross-enterprise data exchange, the data entries and respective devices might still be tampered and therefore build the focal point of this paper. To give an introduction to blockchain devices, current definitions and relevant device configurations, network connections and communication opportunities are gathered through a systematic literature research. The findings are then clustered and discussed with blockchain experts in a semi-structured interview series. Finally, the paper presents a characterization scheme for blockchain devices in form of a multi-dimensional taxonomy and concludes with further research needs. The outcome of the paper also contributes to practice as the taxonomy may also be used as a basis for management decisions.

Keywords

Blockchain Technology; Internet of Things; Industry 4.0; Hardware Devices; Supply Chain Management

1. Introduction

Progressive digitization and globalization lead to various challenges in today's supply chains. Transparency, security and trust are fundamental factors that play a particularly important role in cross-company business processes between both cooperating and competing companies [1]. Traditional approaches for data exchange often fail to manage relevant information in a way that is both transparent for business partners and at the same time safe and trustworthy [2]. Current approaches also pose risks in terms of system failures, integrity, authenticity, and performance bottlenecks [3]. As a result, companies are striving to adapt to these changing conditions and pilot blockchain solutions in various industries [4,5]. Blockchain technology pursues a decentralized approach for data storage and management creating enormous potential for numerous use cases [6]. Logistics and supply chain management in particular pose a suitable application area, as information can be securely exchanged across the entire value chain [7,8]. Still, for this purpose the help of additional technologies, such as devices to access data from the Internet of Things (IoT) is necessary [9]. Many studies on the possibility of using blockchain technology in real world use cases neglect this necessary interplay of technologies. Nevertheless, only by having the right and correct data stored on the blockchain, it makes sense to benefit from its technical functionalities such as immutability and tamper-proof storage [10].

Until today, numerous blockchain projects remained in a Proof-of-Concept (PoC) status as they did not manage to organize a sufficient interplay of technologies and integrate proper devices in their blockchain networks [11,12]. To address this issue, the Ministry of Economic Affairs, Innovation, Digitalization and

DOI: https://doi.org/10.15488/12170

Energy of North Rhine-Westphalia is funding the Blockchain Europe Project and supports this research paper that answers the following research questions:

1. What is a blockchain device and by the use of which dimensions can it be characterized?

2. Which ways exist to integrate devices in a blockchain system and which identity and security mechanisms need to be considered?

To answer these research questions, in the next chapter necessary background information on IoT and blockchain devices is explained. After that the systematic literature procedure, taxonomy development and expert interview approach are explained as used methodologies. Finally, a characterization scheme is presented in form of a taxonomy and discussed in detail. The paper concludes with a summary and further research needs.

2. Background and state of the art

From IoT devices to blockchain devices

An Internet of Things (IoT) device is a physical object that has mechanical or electrical components [13]. It is also "smart" because it is equipped with sensors and microprocessors, enabling the IoT device to perceive and process its environment [14,15]. A further essential characteristic of an IoT device is its digital networking with other devices via standard internet technologies, enabling IoT devices to communicate and perform their tasks automatically [13]. Via equipment systems, such as a monitor, it is possible for humans to interact with the IoT device. Logistics and supply chain management are one of the main application domains for linking blockchain with the Internet of Things and respective devices. Due to the interconnection of resources and goods, both within and across companies, which exchange their states or negotiate interactions, secure storage locations are necessary to keep track of the value-adding activities. Here, the blockchain enables communication between IoT devices as well as the verifiable transmission of information. When used in conjunction with smart contracts, industrial equipment can autonomously provide paid services, report maintenance needs, issue invoices, and make debits. An example of the data that can be exchanged within supply chains is vehicle maintenance data and wear data measured in real time and transmitted directly, conditions such as fill level indicators, derivative indications, or temperature indications for goods subject to a refrigerated container warranty [16]. All of this information triggers follow-up actions, such as intervening when a temperature is exceeded based on tolerance limits or when a vehicle maintenance due date has been exceeded [17].

Blockchain devices - a status quo

Blockchain devices can be represented by different IoT devices, e.g. smartphones, tablets, temperature sensors, or hardware wallets (storage of tokens) that communicate with a blockchain. The first approaches to blockchain devices can be found in the literature: Griggs et al. (2018) designed a blockchain system based on a private Ethereum framework. In this system, sensors communicate with IoT devices that invoke smart contracts and write records of all events on-chain. The IoT device in this system builds a link between sensors and blockchain nodes. The device comes into play as a smartphone that makes patient data visible via appropriate software [18]. Caro et al. (2018) developed a blockchain-based food tracking solution on Ethereum and Hyperledger Sawtooth. In this solution, IoT devices are integrated to process GPS data. By collecting and processing the data directly, IoT devices have direct access to the data and store it as a full node on the blockchain system, ensuring transparent and verifiable traceability. On a truck installed devices scan the batch packaging via an RFID tag and thereby identifiy current goods. When the truck starts moving, the device starts monitoring the temperature and GPS position. [19]. Laszka et al. (2017) describe a privacy preserving energy transactions (PETra) solution for transactive microgrids that allows consumers to trade energy without sacrificing their privacy. PETra is built on distributed ledgers and provides anonymity for communication, bidding, and trading. In this solution, the development of an IoT infrastructure is described, but it is not defined exactly which device communicates with the blockchain and how. The device mentioned is a smart meter, which must be deployed and authorized at each prosumer (producer and consumer) to measure the prosumers' energy production and consumption in a tamper-proof manner [20]. Grecuccio et al. (2020) report a development of a software framework that enables IoT devices to interact directly with an Ethereum-based blockchain. This solution provides an alternative way to integrate a broad category of IoT devices without relying on a centralized intermediary and third-party service. Each IoT device has its own gateway and can sign transactions locally and offline. Moreover, each IoT device is identified by its address within the blockchain and can thus be a target for potential smart contract events [10].

3. Methodology

Structured literature review

Methodically, a systematic literature review according to 21 was conducted for the scientific base of this paper it is highly suitable for opening up emerging topics. We applied numerous search strings in different combinations. "Blockchain Technology" and "Distributed Ledger Technology" in combination with "Devices", "Internet of Things", "Smart Devices", "IoT Devices", "CPS", "CPPS", were used as keywords. In order to narrow down the field of observation in some places, "Supply Chain Management", "Logistics", or "Enterprise Networks" were additionally added to exclude paper without real world application. Finally, we collected 38 relevant papers that we extended by 12 papers through forward and backward research.

Taxonomy development

Following up on the literature review, we developed a taxonomy based on the approach of 22. The method with its roots in Information Systems (IS) research consists of seven steps (see **Figure 1**). First, one must establish a meta-feature that defines the purpose of the taxonomy. Second, end conditions must be established, and an approach must be chosen. The choices are the conceptual-empirical approach or the empirical-conceptual approach. Each approach is divided into three steps. In the conceptual-empirical approach, the focus is on conceptualizing features and dimensions before examining the objects and then creating a taxonomy, while in the empirical-conceptual approach, the focus is on extracting features and dimensions from the objects before grouping them into a taxonomy. In all iterations, we followed the empirical-conceptual approach. These two approaches need to be run repeatedly until the final conditions are met. Nickerson et al. (2013) defined 13 end conditions, divided into eight objective and five subjective (concise, robust, comprehensive, extensible, and explanatory) conditions. We describe the development of our taxonomy and meta-feature, as well as the dimensions with their features, in the next section.

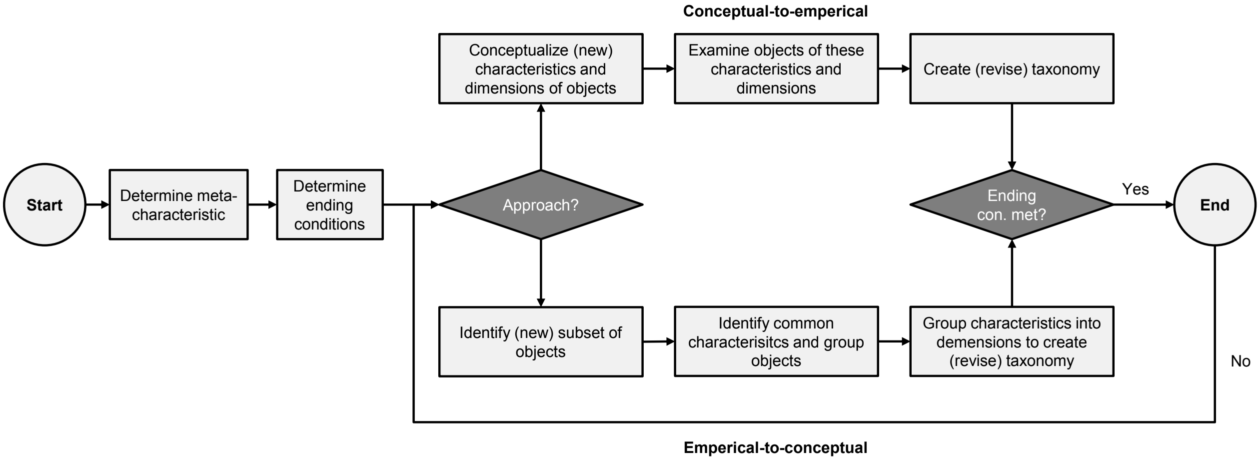


Figure 1: Taxonomy development according to Nickerson et al. (2013)

Expert interviews

The main goal of our taxonomy is to characterize blockchain devices in order to differentiate the large number of devices in the blockchain ecosystem. Therefore, we defined "core features and their feature manifestations of IoT Devices connected to a Blockchain network" as a meta-characteristic for our taxonomy. This meta-characteristic was the basis for identifying additional dimensions and characteristics and did not change during the iterations. The development of the taxonomy required six iterations until we met all 13 final conditions and thus reached the final state. In all iterations, we followed the "Empirical-Conceptual" approach. For additional discussion of the taxonomy with practitioners and to elicit new as well as elaborate on already identified requirements for blockchain devices, semi-structured guided expert interviews were conducted and transcribed according to the clean verbatim transcription approach [23]. The analysis was conducted along the methodology of qualitative content analysis according to 24. The experts listed in Table 1 were approached for interviews.

According to the guidelines, the interviews were scheduled for 45 minutes each. Due to additional explanations, some interviews took a longer time. The backgrounds of the experts cover the fields of business, computer science, logistics, supply chain management and mechanical engineering. The organizations represent application-oriented research institutes and consultancies, as well as a small and large company that both already implemented blockchain applications. All experts have been working on the blockchain topic for at least one year. During each iteration, the derived dimensions for the taxonomy and respective features were discussed, added or deleted.

Table 1: Overview expert interviews

Title	Company	Industry	Date	Duration
Blockchain Developer	Consulting company for enterprise blockchain solutions	Consulting	May 2021	00:45 h
Blockchain Researcher	University chair researching in decentralised markets	Research	May 2021	00:51 h
Consultant	Logistics Service Provider	Logistics	June 2021	01:23 h
CEO	Blockchain Start-up	Logistics and Technology	June 2021	01:16 h
Blockchain Expert	Consulting company for enterprise blockchain solutions	Consulting	June 2021	00:55 h
Researcher in the field of CPPS	University chair researching in CPPS and connection to blockchain systems	Research	June 2021	00:45 h

In the first iteration, we analysed the first 38 papers for the basis of our taxonomy and met seven final conditions set by Nickerson. In this state, we identified 19 dimensions with multiple characteristics each. After discussions in the author team, we realized that some dimensions were not meaningful and in addition duplications occurred, so we decided to reduce five dimensions. In the second and third iteration, we examined the additional 12 papers. After the iterations, we discussed 14 dimensions and met ten final conditions. In the fourth iteration, we decided against categorizing industry use cases and added new dimensions, such as system characteristics, which were used to categorize hardware together with two experts from the interviews conducted in May. Due to ambiguity about whether the system features were low-end or high-end, we decided against it in the fifth iteration after discussions with two further experts from the interviews series in June. We also removed dimensions that related too strongly to IoT. Instead, we added three major layers, which evolved into the final Device Layer, Integration Layer, and Blockchain Layer as we went along. In the last iteration, the taxonomy was finally discussed together in the last two experts from the June interview series and some features were replaced, for example, concrete consensus algorithms emerged instead of categories of consensus algorithms as before, which again was difficult to prove. This resulted in a concise, robust, comprehensive, extensible, and explanatory taxonomy that does not include repetitive dimensions or features to classify all objects identified in the literature review.

4. Findings and discussion

The taxonomy serves as an answer to the research questions of this paper, as it characterizes types of blockchain devices and requirements for a blockchain device to securely and reliably put data on the blockchain. As shown in Table 2, the taxonomy consists of eleven dimensions with 44 characteristics. To increase the transparency and understanding of the taxonomy, we grouped the identified dimensions into the three layers: Device differentiation, blockchain integration, as well as identity and security mechanisms. In addition, an asterisk (*) at the end of a feature indicates whether it is also possible to select multiple characteristics. We visualise the taxonomy as a morphological box, since this is a common way of visualising a taxonomy and it generally illustrates the set of relations contained in a problem complex in an intuitive way [25,26].

Device differentiation

Performance: IoT devices can basically be divided into low-end and high-end on the criteria of their hardware equipment (computing power, storage capacity, battery capacity, communication capability, etc.). The main difference is the executability of software based on traditional operating systems (such as Linux). While a low-end device cannot execute software based on traditional operating systems, the high-end device is able to execute such software [27–29]. In addition, limited memory capacity, communication capability (broadband), and computational power are not particularly suitable to support resource-intensive distributed ledgers, but sufficient to be able to use the services of a distributed ledger network (e.g., by using an API) [29].

Equipped systems: The interaction with the environment functions via measurement, control and regulation technology. Only sensors that perceive the environment are used for measurement technology. For example, temperature sensors perceive the outside temperature or pressure sensors perceive certain forces [15]. In contrast, control technology exclusively uses actuators. Actuators are the signal converter counterpart to sensors and form the actuators in a control loop, i.e. they convert signals (e.g. commands from the control computer) into mechanical movement or other physical variables (e.g. pressure or temperature). In control technology both, sensors and actuators, are used. The sensor system displays current measured values, while the actuator system triggers a specific action when a measured value is exceeded [18,27,30].

Communication Technologies: One of the main features of blockchain devices is their communication with each other or with other systems via internet. Communication technologies include wired technologies such as the Local Area Network (Ethernet, PLC, bus systems) and wireless ones, such as Wireless Personal Area Networks that represent the most widely used communication technology among IoT devices and have a range of approximately 100m. Examples of WPAN are devices with Bluetooth, ZigBee or Z-Wave equipment [32,29,31]. WLAN, as another wireless communication technology and enables ranges of up to 1 km. In this category, Wi-Fi is the most widely used standard [29]. LPWAN is a communication technology that is predicted to grow rapidly. Key factors are the extremely long battery life and a maximum communication range of over 20 km [29]. Mobile networks, such as GSM, 3G, 4G and 5G, are used for the long-range operation of IoT devices. 2G, 3G and 4G technologies have long been the only option for device connectivity. Now that LPWAN and also 5G are gaining prominence, these legacy mobile standards are expected to give up their share to the new technologies [29,31].

Table 2: Taxonomy for Blockchain Devices

MD	Dimensions	Characteristics							MEX
Device Differentiation	Performance	Low-end Device	High-end Device						Y
	Equipped Systems	Sensor System	Actuator System	Control System	Regulation System	Visual Indication			N
	Communication Technologies	Wired	Wireless						N
		Local Area Network	Sigfox	Software Defined Networks	Personal Area Network	Wireless Area Network	Mobile	Neul	N
Network Integration	IT Architecture	Centralized	Decentralized						Y
	Network Topology	Star	Point-to-Point	Mesh					Y
	Blockchain Governance	Independent Blockchain-Network	Participating (Connection to existing Blockchain-Network)	Integrated (BaaS, Cloud-based)					Y
	Blockchain Types	Public Permissionless	Consortium	Private Permissioned					Y
	Blockchain Identifiability	No Node	Light(weight)-Node	Full-Node					Y
	Gateway	Cloud Server	Enterprise Server	Other Device	No Gateway				Y
Identification and Security	Identity Management	Self-Sovereign Identity	Bring Your Own Identity	Public Key Infrastructure	Decentralized Public Key Infrastructure				Y
	Security Mechanism	Anonym Digital Signatures	Non-interactive Zero-Kowledge Proofs	Homomorphic Encryption Algorithm	Secure Multiparty Calculation Protocol	Attribute-based Encryption Algorithm	Mixing Procedures		N

MD = Meta-Dimensions; MEX = Exclusivity

Network integration

IT architecture: IT architectures especially in the area of IoT can be divided into centralized and a decentralized types. In centralized architectures, a central hub is used to provide backend services for smart devices. Some of the most important centralized capabilities are event processing, events notification and real-time analytics. In addition to the mentioned capabilities there are also scenarios where decentralized communication between IoT devices is required without the need of a central hub. There are many examples of decentralized IoT applications like peer-to-peer messaging or decentralized auditing and file sharing. [4,31]

Network topologies: The network topologies which are used in IoT can be split into three categories: star, point-to-point and mesh. The point-to-point topology is based on a direct connection between the nodes. In star networks every device is connected to a central hub. In a mesh network topology every node can be connected with each other. There are six networking attributes: latency, throughput, fault resiliency, scalability, the number of hops and range. These attributes can help developers of IoT Applications in knowing the capabilities of the different network topologies to choosing the best topology for their own Application [31].

Blockchain-Governance: Blockchain governance determines the organisational structure, jurisdictions in and requirements for the usage of blockchain-based applications as well as the consortium agreement process. Three governance structures can be distinguished. On the one hand, an independent blockchain network can be established and managed for the individual use case of a company. On the other hand, it is possible to join and participate in already existing blockchain networks and accustom to the already existing governance. Finally, it is also possible to make use of external service providers who make a blockchain network available. These include blockchain-as-a-service- or cloud-based solutions. [33,34]

Blockchain types: public blockchain represents an ecosystem, publicly visible to everyone. This type of blockchain has become known through the crypto networks Bitcoin and Ethereum. [35] Private blockchains (e.g. Multichain) offer governance rules that have to be developed individually during network construction.

With this type, data is shared in a restrictive manner and participants can only view defined transactions. A consortium blockchain is a special form of private blockchain because the consensus process participation is distributed among several organisations in the P2P network. The transaction activity is isolated from the public [35]. Hyperledger Framework of the Linux Foundation can be mentioned as an exemplary framework. A hybrid blockchain offers public access to the network for everyone and at the same time a trust-based governance structure. [36]

Blockchain identifiability: IoT Devices without a direct connection to the blockchain network cannot be identified via the blockchain, as they communicate indirectly (e.g. via a cloud server) with the blockchain network. [10] However, there are also devices that can identify themselves in a blockchain network. On the on hand, there are light clients which are nodes with a computing capacity and network bandwidth that is too low to download and check the entire blockchain [37]. Ethereum has a client application named Mist Browser, a user-friendly wallet also known as a Light Node. This Light Node connects to a blockchain to perform only basic functions of a full node, such as sending and receiving cryptocurrencies which only requires a wallet application on the IoT device [38,39]. A light node is thus able to sign and broadcast transactions on its own. On the other hand, there are devices that hold a full copy of the blockchain and have sufficient processing and storage capacity to act as a Full Node. IoT home gateways, for example a Raspberry PI, can already participate in the blockchain as a full node and thus potentially support blockchains [29].

Gateway: One possible solution for connecting devices consists of a communication between a central cloud server and the devices, also called IoT cloud server. The server is responsible for collecting data from the IoT devices and for storing this data in the blockchain. One of the weak points of this solution is the central server as a single point of failure. Another crucial vulnerability is the lack of digital signatures of the IoT devices. The data that is sent to the blockchain is not signed on the spot by the device, but only when the data is received by the central server. This means that the authenticity and integrity cannot be guaranteed from the source [10]. Based on the remote procedure call (RPC) developed by Birrell and Nelson in 1984 [40], there is the possibility of triggering the execution of a procedure on a remote enterprise server through embedded gateways. The enterprise server provides the gateways with an API, which enables interaction with the blockchain. The gateways should be uniquely identifiable, sign transactions locally and offline with their private keys before communicating with the RPC server. Using its own addresses within the blockchain, each IoT device can be identified and thus be a target for possible smart contract events [10]. Apart from that, other devices can act as gateways to enable a communication from smaller low-end-devices to the blockchain [37].

Identity and security management

Identity management: A relatively new approach to identity management is the Self-Sovereign Identity (SSI) paradigm based on decentralized infrastructures. Typical for SSI is the focus on the user of the digital identity, who is in possession of his personal data himself and decides on third-party access [41]. The user receives identity features and corrections in the form of cryptographically secured digital proofs - the verifiable credentials - and can manage them independently by means of a digital wallet [42]. Bring Your Own Identity (BYOI) in this context refers to the idea and goal of being able to use this own identity on demand in any environment, be it private or business [41]. Public Key Infrastructures (PKI) are one of the mechanisms for managing keys in public key cryptographic systems. A private key owned only by the user allows the signing of different contents and documents. The public key then allows anyone to verify the respective signature [43]. Efforts to adapt PKI systems to emerging challenges, result in the development of Decentralized PKI (DPKI). One guiding and already practiced idea is the hierarchy-free web-of-trust, in which users mutually confirm credibility and correctness of associated data and trust in an assigned public key while network partners authenticating it [41].

Security mechanisms: The idea of anonymous digital signatures is that users or objects within the blockchain network can use pseudonyms to hide their true identity and thus secure their privacy [14,44]. Mixing procedures involve mixing users' or objects' values with each other, which leads to confusion within the network. The identities can be disguised with this mechanism. To protect digital assets from attackers, Mixcoin, for example, obfuscates users by mixing currencies simultaneously and also uses an accountability mechanism to detect asset thefts [14,44]. The homomorphic encryption algorithm is a technique that enables computations to be performed on the cipher text itself. Hence, it is not necessary to convert data into plaintext in order to perform an operation on it. Homomorphic cryptography can be easily applied to the data on-chain without changing the blockchain properties, which ensures privacy and allows data to be verified and managed only in encrypted form. Secure Multiparty Computation Protocols (SMCP) are a class of algorithms that allow a group of mutually untrusting actors to evaluate functions without having to reveal their private inputs [14,44]. Attribute Based Encryption (ABE) is a cryptographic algorithm that uses the attributes as regulatory factors for the cipher text encrypted with the user's private key. The text data can only be decrypted if the attributes of the decoders match the encrypted data [14,44].

5. Conclusion

Previous research has provided good reasons to believe that blockchain solutions will diffuse in various industries over time. To exploit all functionalities of the technologies, getting the right data in an integer and traceable manner on chain, constitutes an important challenge. Appropriate configured blockchain devices address this challenge and are described in this paper by (1) hardware constitution as well as (2) technical possibilities to connect them to the blockchain systems and (3) operate identity as well as security management measures.

The central outcome of the paper is a taxonomy characterizing relevant dimensions of blockchain devices, scientifically substantiated by literature research and expert interviews. It became clear that current enterprise blockchain projects use different blockchain devices with different and individual characteristics. Thus, a strict determination of a singular blockchain device is not possible. However, our characteristics help to understand the range of different device types and possibilities to integrate them in a blockchain system. Therefore, in a first dimension, the devices are differentiated according to their performance and energy efficiency as well as equipped system and communication technologies. A second dimension addresses the type and topology of IoT platform the device gains access to; the type of governance and framework of the connected blockchain system; as well as the way the device is identified by and connected with the system. A final layer addresses possible identity and security management measures for the device.

This outcome aims to advance previous research on devices that are described in blockchain research and delivers a first-ever possibility to classify blockchain devices. The investigation is of considerable relevance to blockchain scholars as well as practitioners that find themselves in PoC blockchain projects and work on device integrations. In order to further validate the developed taxonomy real world blockchain devices will be described by its means in a future research work by the Blockchain Europe Project. Blockchain scholars are invited to build up on the taxonomy and apply it in further case studies to demonstrate international acceptance and applicability in practice.

Acknowledgments

The work was funded by the Ministry of Economic Affairs, Innovation, Digitalization and Energy of the State of North Rhine-Westphalia.

References

[1] Henke, M., 2003. Strategische Kooperationen im Mittelstand: Potentiale des Coopetition-Konzeptes für kleine und mittlere Unternehmen (KMU). Zugl.: München, Techn. Univ., Diss., 2002 u.d.T.: Henke, Michael: Strategische Kooperationen kleinerer und mittlerer Unternehmen (KMU) unter besonderer Berücksichtigung des Coopetitions-Ansatzes. Verl. Wiss. & Praxis, Sternenfels, 208 pp.

[2] Gelhaar, J., Guerpinar, T., Henke, M., Otto, B., 2021. Towards a Taxonomy of Incentive Mechanisms for Data Sharing in Data Ecosystems. Asia Conference on Information Systems.

[3] Neugebauer, R. (Ed.), 2018. Digitalisierung, 1. Aufl. 2018 ed. Springer Berlin Heidelberg, 416 pp.

[4] Große, N., Leisen, D., Gürpinar, T., Forsthövel, R.S., Henke, M., ten Hompel, 2020. Evaluation of (De-) Centralized IT technologies in the fields of Cyber-Physical Production Systems. CPSL.

[5] Gürpinar, T., Guadiana, G., Ioannidis, P.A., Straub, N., Henke, M., 2021. The Current State of Blockchain Applications in Supply Chain Management, 168–175.

[6] Heines, R., Gürpinar, T., 2021. Towards a Typology of Blockchain-based Applications: a Conceptualization from a Business Perspective. BAS21 Mittweida.

[7] Grosse, N., Guerpinar, T., Henke, M., 2021. Blockchain-Enabled Trust in Intercompany Networks Applying the Agency Theory, in: 2021 3rd Blockchain and Internet of Things Conference. Ho Chi Minh City Vietnam.

[8] Gürpinar, T., Harre, S., Henke, M., Saleh, F., 2020. Blockchain Technology – Integration in Supply Chain Processes. Hamburg International Conference of Logistics.

[9] Dujak, D., Sajter, D., 2019. Blockchain Applications in Supply Chain, in: Kawa, A., Maryniak, A. (Eds.), SMART Supply Network. Springer International Publishing, Cham, pp. 21–46.

[10]Grecuccio, J., Giusto, E., Fiori, F., Rebaudengo, M., 2020. Combining Blockchain and IoT: Food-Chain Traceability and Beyond. Energies 13 (15), 3820.

[11]Mika, B., Goudz, A., 2019. Blockchain-Technologie in der Energiewirtschaft: Blockchain als Treiber der Energiewende, 1. Aufl. 2020 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 113 pp.

[12]Varriale, S., 2019. Unequal Youth Migrations: Exploring the Synchrony between Social Ageing and Social Mobility among Post-Crisis European Migrants. Sociology 53 (6), 1160–1176.

[13]Hompel, M. ten, Bauernhansl, T., Vogel-Heuser, B., 2020. Handbuch Industrie 4.0. Springer Berlin Heidelberg, Berlin, Heidelberg, 631 pp.

[14]Fridgen, G., Guggenberger, N., Hoeren, T., Prinz, W., Urbach, N., 2019. Chancen und Herausforderungen von DLT (Blockchain) in Mobilität und Logistik. https://www.bmvi.de/SharedDocs/DE/Anlage/DG/blockchain-gutachten.pdf?__blob=publicationFile.

[15]Kruse Brandão, T., Wolfram, G. (Eds.), 2018. Digital Connection. Springer Fachmedien Wiesbaden, Wiesbaden.

[16]Pandey, R., 2001. Essentials of Supply Chain Management. OPSEARCH 38 (2), 238–239.

[17]Voß, P.H., 2020. Logistik – die unterschätzte Zukunftsindustrie. Springer Fachmedien Wiesbaden, Wiesbaden.

[18]Griggs, K.N., Ossipova, O., Kohlios, C.P., Baccarini, A.N., Howson, E.A., Hayajneh, T., 2018. Healthcare Blockchain System Using Smart Contracts for Secure Automated Remote Patient Monitoring. Journal of medical systems 42 (7), 130.

[19]Caro, M.P., Ali, M.S., Vecchio, M., Giaffreda, R., 2018. Blockchain-based traceability in Agri-Food supply chain management: A practical implementation, 1–4.

[20]Laszka, A., Dubey, A., Walker, M., Schmidt, D., 2017. Providing privacy, safety, and security in IoT-based transactive energy systems using distributed ledgers, 1–8.

[21]Baumeister, R.F., Leary, M.R., 1997. Writing Narrative Literature Reviews. Review of General Psychology 1 (3), 311–320.

[22]Nickerson, R.C., Varshney, U., Muntermann, J., 2013. A method for taxonomy development and its application in information systems. European Journal of Information Systems 22 (3), 336–359.

[23]McLellan, E., MacQueen, K.M., Neidig, J.L., 2003. Beyond the Qualitative Interview: Data Preparation and Transcription. Field Methods 15 (1), 63–84.

[24]Mayring, P., Fenzl, T., 2019. Qualitative Inhaltsanalyse, in: Baur, N., Blasius, J. (Eds.), Handbuch Methoden der empirischen Sozialforschung. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 633–648.

[25]Ritchey, T., 2006. Problem structuring using computer-aided morphological analysis. Journal of the Operational Research Society 57 (7), 792–801.

[26]Szopinski, D., Schoormann, T., and Kundisch, D., 2020. Visualize Different: Towards Researching the Fit Between Taxonomy Visualizations and Taxonomy Tasks. Proceedings of the Wirtschaftsinformatik.

[27]Hahm, O., Baccelli, E., Petersen, H., Tsiftes, N., 2016. Operating Systems for Low-End Devices in the Internet of Things: A Survey. IEEE Internet Things J. 3 (5), 720–734.

[28]Noura, M., Atiquzzaman, M., Gaedke, M., 2019. Interoperability in Internet of Things: Taxonomies and Open Challenges. Mobile Netw Appl 24 (3), 796–809.

[29]Vermesan, O., 2018. Next Generation Internet of Things. River Publishers, Aalborg, 352 pp.

[30]Hasenjäger, E., 2015. Regelungstechnik für Dummies: Auf einen Blick: Prozesse, Reglertypen, Regelkreise und die mathematischen Gleichungen verstehen, 1. Aufl. ed. Wiley-VCH, Weinheim, 443 pp.

[31]Yaqoob, I., Ahmed, E., Hashem, I.A.T., Ahmed, A.I.A., Gani, A., Imran, M., Guizani, M., 2017. Internet of Things Architecture: Recent Advances, Taxonomy, Requirements, and Open Challenges. IEEE Wireless Commun. 24 (3), 10–16.

[32]Hinckeldeyn, J., 2019. Blockchain-Technologie in der Supply Chain: Einführung und Anwendungsbeispiele. Springer Fachmedien Wiesbaden; Imprint: Springer Vieweg, Wiesbaden, 56).

[33]van Pelt, R., Jansen, S., Baars, D., Overbeek, S., 2021. Defining Blockchain Governance: A Framework for Analysis and Comparison. Information Systems Management 38 (1), 21–41.

[34]Werner, J., Zarnekow, R., 2020. Governance of Blockchain-Based Platforms, in: Gronau, N., Heine, M., Krasnova, H., Pousttchi, K. (Eds.), Proceedings der 15. Internationalen Tagung Wirtschaftsinformatik 2020. GITO mbH Verlag für Industrielle Informationstechnik und Organisation, Berlin, pp. 128–141.

[35]Hileman, G., Rauchs, M., 2017. 2017 Global Blockchain Benchmarking Study. SSRN Journal.

[36]Tobin, Reed, 2017. The Inevitable Rise of Self Sovereign Identity. Sovrin Foundation. https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf. Accessed 27 July 2021.

[37]Al-Bassam, M., Sonnino, A., Buterin, V., 2018. Fraud and Data Availability Proofs: Maximising Light Client Security and Scaling Blockchains with Dishonest Majorities, 33 pp. http://arxiv.org/pdf/1809.09044v5.

[38]Dannen, C., 2017. Introducing Ethereum and Solidity. Apress, Berkeley, CA.

[39]Gürpinar, T., Ashraf, S.R.B., Broza-Abut, N., Sparer, D., 2022. Blockchain-Based Infrastructure for Product Traceability in the Medical Supply Chain, in: Moumtzoglou, A., Borah, M.D., Zhang, P., Deka, G.C. (Eds.), Prospects of Blockchain Technology for Accelerating Scientific Advancement in Healthcare. IGI Global, pp. 119–134.

[40]Birrell, A.D., Nelson, B.J., 1984. Implementing remote procedure calls. ACM Trans. Comput. Syst. 2 (1), 39–59.

[41]DIN SPEC 3103:2019-06, 2019. Blockchain und Distributed Ledger Technologien in Anwendungsszenarien für Industrie 4.0.

[42]Ehrlich, T., Richter, D., Meisel, M., Anke, J., 2021. Self-Sovereign Identity als Grundlage für universell einsetzbare digitale Identitäten. HMD 58 (2), 247–270.

[43]Pal, O., Alam, B., Thakur, V., Singh, S., 2021. Key management for blockchain technology. ICT Express 7, 76–80.

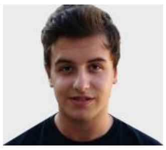
[44] Idrees, S.M., Nowostawski, M., Jameel, R., Mourya, A.K., 2021. Security Aspects of Blockchain Technology Intended for Industrial Applications. Electronics 10 (8), 951.

Biography

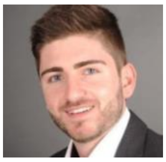
Tan Gürpinar, M.Sc. (*1991) is a researcher and field coordinator for blockchain technology at the Chair of Enterprise Logistics, TU Dortmund University. His work focuses on the effective and profitable integration of new technologies, especially blockchain technologies.

Dr.-Ing. Max Austerjost (*1986) is a researcher at Fraunhofer IML. He leads Blockchain Europe, the project to establish the European Blockchain Institute in NRW. His special research interest is making blockchain technology usable in supply chain management and logistics.

Dipl.-Ing. Josef Kamphues (*1986) is Head of Department Supply Chain Development & Strategy at Fraunhofer IML. His work focuses on the research and application of digital technologies in supply chain management with a focus on blockchain, AI and simulation.



Jonas Maaßen, B.Sc. (*1997) studies Logistics at TU Dortmund University and works as a research assistant for blockchain technology at the Chair of Enterprise Logistics, TU Dortmund University.



Furkan Yildirim, M.Sc. (*1992) is a researcher for blockchain technology at the Chair of Enterprise Logistics, TU Dortmund University. His research focuses on the utilization of tokenomics in supply chain management.

Univ.-Prof. Dr. habil. Michael Henke (*1971) is the chairholder of the Chair of Enterprise Logistics, TU Dortmund University and director of the Fraunhofer Institute of Material Flow and Logistics since 2013.

3rd Conference on Production Systems and Logistics

Towards A Flexible Approach To Transfer Machine Operation Know-How From Experts To Beginners With AI

Timo Leitritz[1], Martina Köhler[1], Christian Jauch[1]
[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany*

Abstract

Training new users at a production machine is a time intensive and expensive task. To reduce the effort in this task we examine the possibilities of enhancing a production machine with a system that is able to learn from its users and teach inexperienced users this knowledge: Self-Learning and Self-Explanatory Machine SLEM. The learning process of SLEM relies on watching an experienced user working on a machine using camera-based human activity recognition which predicts the acitivities based on the estimated human skeleton in the video stream. SLEM must be able to work with little data to reduce the learning time as much as possible. Thus, this paper shows that training an activity recognition model solely on one experienced individual's actions can lead to comparatively high activity recognition accuracy despite the low data variety. The results show that training on a single-person dataset can reach relatively high accuracy levels and is a suitable way of training the model in the industrial setting. For the teaching process, in which the system has to compare the actual activities with the target acitivities to give feedback, the activity recognition has to run in real-time. Different amounts of input data for the activity recognition model are examined and lead to a configuration with little accuracy loss and sufficient latency performance.

Keywords

Machine Learning; Human Activity Recognition; Pose Estimation; Skeleton; Industrial; Online Activity Recognition

1. Introduction

Highly complex machines in manufacturing companies need to be operated and maintained by human experts. Furthermore, a lot of older machines are expensive to upgrade even though a lot of the older knowledge about the machine is missing, so they require specialized knowledge from experts. This poses challenges whenever a machine expert leaves the company and is not able to transfer their knowledge to more inexperienced workers. Growing employee turnover causes further important knowledge to get lost. Additionally, teaching inexperienced workers is not only a time intensive task, but also causes considerable machine downtime.

The developed system from the publicly funded project "Self-Learning and Self-Explanatory Machine" (SLEM) tried to solve this issue using a variety of AI-based methods. A camera-based system was able to learn the manual interaction with the machine from an experienced user (subsequently referred to as "experts"). Furthermore, the system was able to generate a guideline for optimal machine operation from this gained knowledge to help inexperienced users ("beginners").

DOI: https://doi.org/10.15488/12127

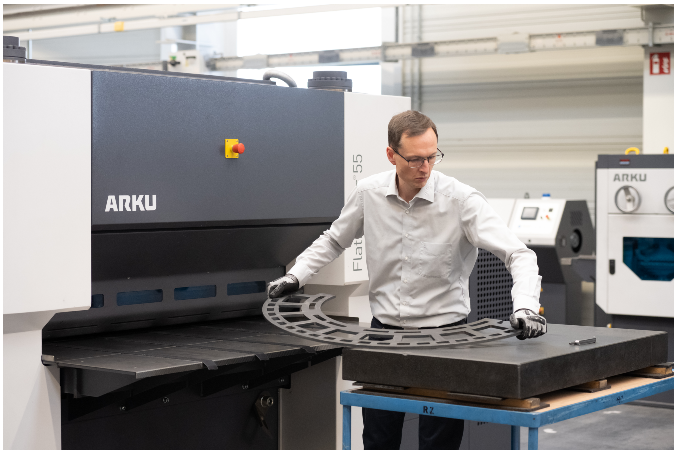


Figure 1: ARKU leveler FlatMaster® 55

This paper presents the results of the SLEM project on skeleton based activity recognition using PoseC3D [1], as well as comparing two types of scenario driven datasets: "Expert" and "Mixed". The former contains data only from one expert user, whereas the latter combines data from a variety of differently skilled users. A skeleton based Human Activity Recognition (SHAR) was chosen due to the highly generalized options for pose recognition that are available as well as due to the ability to abstract from the image based representation of a human, that can differ greatly regarding the appearance, e.g. in cloth type or hair color. The following topics are addressed in this paper:

- Recording, annotating and preprocessing a scenario driven dataset from a real world use case at a company, see Figure 1
- Integrate PoseC3D in an application oriented real time activity recognition pipeline for an actual industrial use case
- Examination of accuracy on "Expert" and "Mixed" dataset
- Examination of different input clip sizes and pose estimation intervals to optimize system latency

2. Related work

Human Activity Recognition (HAR) in industrial settings promises many benefits for assistance systems in manual work. An AI-based assistance system with implemented HAR is able to automatically detect and predict what the human is currently doing or even going to do. Based on this information the assistance system is able to react, e.g. by giving support or security warnings. The required activities must first be learned by the HAR method on annotated data and the granularity of the defined activities can vary greatly, from detecting "walking" and "standing still" to precise gestures. In the recent years, many new methods for HAR have been developed, using a variety of different deep learning architectures such as 2D-CNN, 3D-CNN, RNN, GCN and Transformers. In general, these approaches can be further differentiated by the type of input data they use:

- Integrated sensors, e.g. accelerometer data [2] [3]
- Image and video data [4] [5] [6]
- Skeleton data [1] [7] [8]

Data from integrated sensors, such as accelerometer data, is a common data source for HAR since it is available in every smartphone and is a rather inexpensive technology. The detail of activities detected based on such data is limited and mainly focuses on detecting specific movements of the person wearing the sensor. Using image and video data offers a similar data source as a human uses to recognize activities and scenes with the sense of sight. Thus a higher detail of activities can be recognized with the cost of significantly more computation in training and execution of the HAR approach. In addition, a great amount of diverse annotated data is required to train a generalized model and being able to recognize activities on people with varying appearances. In contrast to that, using skeleton data for HAR solves this issue by abstracting from the visual appearance of the person and focusing on keypoints of the human skeleton.

Furthermore, different types of input data can be combined, which was shown in [9], where a unified framework using visual and skeleton features was able to recognize human activities.

Several publications also focused on industrial settings [10] [11] [12]. An end-to-end approach using only skeleton data as input was developed by [10]. A video of an industrial scenario was given and the output was an activity class. First, skeleton data was predicted for each frame in the video through pose estimation, specifically using the stacked hourglass model. Afterwards, leveraging a spatial transformer (STN), the skeleton was further augmented towards translation, rotation, and scaling in order to give the data more human body pose diversity. A graph convolutional network (GCN) [7] was used in the end in order to preserve the relationship between the limbs from which activity scores were calculated. Due to the nature of the augmentation by the STN, the authors found that some activities – such as walking back and forth – are not rotation invariant and disadvantaged the HAR unnecessarily.

A slightly different industry application scenario is the interaction with robots. Since many industrial robots work as human assistances, it becomes more and more important for them to know what the human is doing. [11] used 3D skeleton data and a random forests binary classifier in order to classify primitive movement actions in real time, since such actions could already assist in human-robot interaction.

Another publication focused on using HAR as a way to make interactions between users and industrial machines more natural [12]. The authors used images as raw input for the networks and tested various convolutional networks especially in regards to their execution time. This is particularly important in order to fit into their Natural Machine Operation (NaMO) framework that was developed in order to create a fluid workflow between human and machine and to fit HAR into this workflow. The authors show that it is possible to train HAR models on small industrial datasets.

In general, HAR poses many challenges in industrial settings. Overfitting on small datasets, as well as the effort of having to manually annotate these datasets. The latter poses special challenges when there are no commonly agreed upon classes in such individual cases. However, by using a skeleton-based approach the amount of data necessary to train a generalized model can be reduced, since the skeleton data provides an abstraction of appearance. Thus, this paper focuses on training HAR with limited datasets in terms of variety by only using data from a single person and evaluating the result on two separate individuals, which was not examined in detail by the above publications [10] [11] [12].

3. Methods

Section 3.1 explains the core deep learning pipeline in terms of the model architecture as well as the training and inference pipeline to provide a detailed understanding of the used HAR approach. The data acquisition and preparation techniques are explained in Section 3.2, focusing on all relevant steps to prepare the dataset for training. This is especially important since training and inference were conducted on a custom industrial dataset with custom activity classes. Section 3.3 describes the used metrics in the following experiments to prepare for the evaluation in Chapter 4.

3.1 Deep learning pipeline

The pipeline from video stream to classified activity regarding SHAR consists of two main parts: Pose extraction and activity recognition. For activity recognition PoseC3D [1] was selected as promising SHAR architecture as it achieved the highest accuracy on the NTU60 cross-view test dataset as of January 2022

[13]. It was implemented in the widely used MMAction2 framework by OpenMMLab [14]. PoseC3D uses a 3D-CNN to recognize activities from a 3D heatmap volume, generated from the skeleton data of the video. The key is to convert 2D pose information to a heatmap that encodes the skeleton information of a single frame and can be processed efficiently by a CNN architecture. Each pose keypoint (e.g. elbow, eye, knee) is represented with a Gaussian distribution at the corresponding location as one channel in the resulting heatmap. These heatmaps are generated per frame of the input video clip and then stacked together to a 3D-Heatmap that represents the pose information of the video clip. For pose extraction the same two stage approach was used as in [1]. This approach consists of a Faster R-CNN [15] object detector to detect all people in the image and a HRNet [16] pose estimator to extract the skeleton information per person. The object detector is implemented in the MMDetection framework [17] whereas the pose estimation model is implemented in the MMPose framework [18]. The pose estimator is trained on the COCO [19] body keypoints dataset.

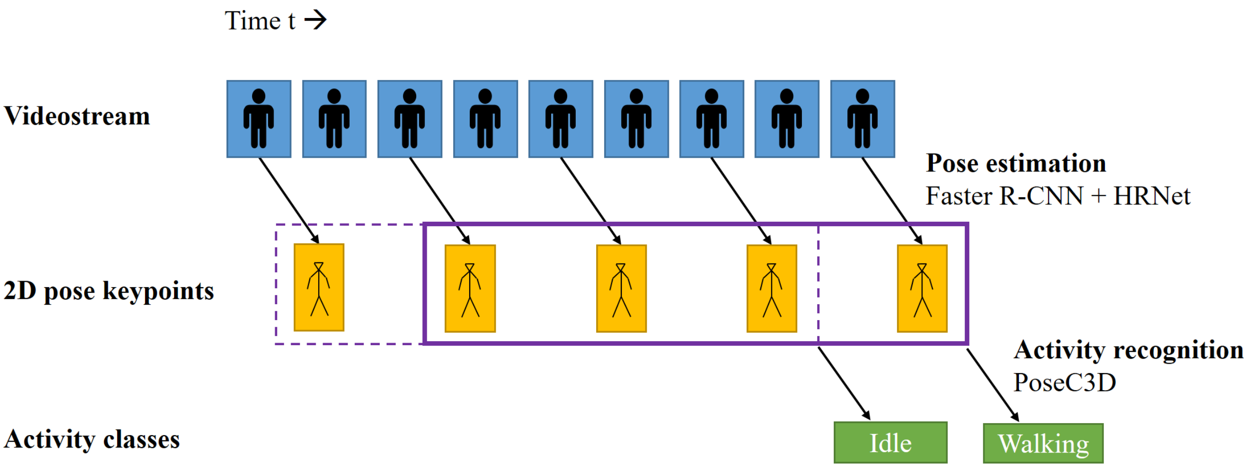


Figure 2: Activity recognition pipeline illustration. The pose estimation is only applied on every second frame, which is due to the pose estimation interval $I_p = 2$. The violet rectangle represents the sliding window of pose data that acts as input for the activity recognizer PoseC3D.

Figure 2 illustrates the SHAR pipeline in an online scenario. In contrast to the NTU60 offline scenario, PoseC3D has to classify video clips of constant length instead of varying length. This is why "Uniform Sampling" as proposed in [1] was not applied in this work. In the online pipeline shown in this paper, PoseC3D takes a fixed amount of N_{PC3D} poses including the current and $N_{PC3D} - 1$ prior poses as inputs. The pose estimation can be executed on each frame to extract the poses, but with a given framerate of 30 fps, this is very computationally expensive. Alternatively, poses can be estimated once every p frames in a pose estimation interval I_p. Estimated poses are saved in a ring buffer of length N_{PC3D} to be partially reused for the next activity recognition execution.

3.2 Data acquisition and preparation

To acquire data that can be used to train and test these models, videos of workers of different skill levels (beginner, intermediate, expert) during usual tasks on a metal leveler machine were recorded (see Figure 1). This was done as part of a Data Driven User Needs Assessment [20], which aims to analyze user experiences of the workspace based on machine data, visual data, eye tracking as well as interviews with the machine operator. Roughly 2 h 40 min (6 expert, 2 intermediate and 3 beginner videos) of 2D-RGB video material in total was recorded with a video resolution of 2560 by 1440 pixels using a Microsoft Kinect Azure DK camera sensor. In an offline post-processing step the 2D pose data of each frame is extracted from all videos by the pose estimation method described in Section 3.1.

The videos were annotated frame-by-frame, where each frame received exactly one class label. The analysis of the video data resulted in the decision on nine activity classes for working on the leveler machine: *"0 - Idle", "1 - Walk", "2 - Measure with tool", "3 - Rearrange part", "4 - Test part with hands", "5 - Pick object", "6 - Place object", "7 - Interact with machine interface", "8 - Feed part to machine".*

The annotated videos recordings were split into a training and validation dataset with two variants (see Table 1) and one test dataset according to Table 1. In addition, the training and validation dataset was split into training (90%) and validation dataset (10%) after sampling and balancing the dataset. The goal of the sampling was to convert the dataset of large videos with mixed activity annotations into short video clips of fixed size and exactly one activity class. From each video recording, video clips of a fixed clip size were sampled by moving a sliding window with a step size of one over the recording, resembling the online pipeline scenario from Section 3.1. Resulting clips with annotations from more than one activity class were discarded. Furthermore, the balancing was applied due to large differences in the class occurrences in the dataset. A new data set with balanced class occurrences was derived from the sample of clips. The resulting dataset only contained pose data and no image data since PoseC3D only requires pose data.

Table 1: Overview of the amount of video scenarios in training, validation and test datasets

Dataset name	Expert	Intermediate	Beginner	Total
Training/validation "Mixed"	5	1	2	8
Training/validation "Expert"	5	0	0	5
Test	1	1	1	3

3.3 Evaluation metrics

For evaluation and comparison, the accuracy of the resulting predictions was calculated. More specifically, top-K accuracy was used with $K = 1$ and $K = 5$. This means that a prediction was classified as correct, if the actual class occurs in the K highest confidence predictions.

Furthermore, a distinction between "total" top-1 accuracy as well as scenario specific accuracies "Beginner", "Intermediate" and "Expert" was added. The "total" top-1 accuracy is calculated on the whole test dataset, whereas the scenario accuracies only took into account predictions from the corresponding scenario.

In order to evaluate the runtime performance of the activity recognition pipeline, "PoseC3D" latency and "total" latency were measured. The former measured the inference time for a single prediction of the PoseC3D model. The latter took into account the amount of time necessary to run the pose estimation as well as the activity recognition for the specified pose interval and clip input size over a time period of 1.6 s of a video stream. This equals 48 frames at 30 fps.

4. Experiments

This section describes the conducted experiments that were used to examine various aspects of SHAR. The first experiment focused on the performance difference of PoseC3D on the two industrial dataset variants "Expert" and "Mixed" to investigate the impact of dataset with limited variety. Afterwards, an experiment that compared several model input configurations was conducted with different pose estimation intervals and input clip sizes to analyze the trade-off between accuracy and latency for a real-time HAR scenario.

4.1 PoseC3D performance on "Expert" vs. "Mixed" dataset variants

PoseC3D achieves state-of-the-art accuracy on the common and publicly available dataset NTU60 [21]. This experiment aims to show the performance of PoseC3D on the custom industrial dataset. Furthermore, it was simultaneously examined how well the model can learn from the specialized “Expert” dataset variant in contrast to training on the “Mixed” dataset. It was assumed that models trained on purely the “Expert” dataset would have issues generalizing towards the “Beginner” and “Intermediate” videos.

Training hyper parameters were similar to the default MMAction2 configuration for PoseC3D: Learning rate was set to 0.01 with SGD and batch size adjusted to 16. All clips in the dataset had a length of 48 pose frames. PoseC3D’s clip input size was set to 48, so no “Uniform Sampling” was used as the frame number matches the dataset video clip size. A starting checkpoint of PoseC3D that was pre-trained on the NTU60 dataset was used. The training ran on a single NVIDIA Tesla V100 GPU with no augmentation applied. The random crop and flip operations were explicitly removed in the PoseC3D data pipeline, in order to create a baseline without any augmentation that can be built upon in future experiments.

Table 2: PoseC3D accuracy results after training on “Mixed” and “Expert” dataset for 10 epochs with $I_p = 1$ and $N_{PC3D} = 48$

	Top-1 accuracy				Top-5 accuracy
Model name	Total	Beginner	Intermediate	Expert	Total
PC3D-48-1 „Mixed“	0.456	**0.554**	0.418	0.449	0.910
PC3D-48-1 „Expert“	**0.508**	0.433	**0.487**	**0.584**	**0.952**

The results of the evaluation in Table 2 show that PoseC3D did not reach the same accuracy level as on NTU60, where it achieved a cross-subject accuracy of 0.941. The best total accuracy result occurred when training on the “Expert” dataset with a total top-1 accuracy of 0.508 whereas the training on the “Mixed” dataset resulted in a drop of 10.2 %. Top-5 accuracy is close between both variants. Since the test dataset includes different amounts of samples per scenario (312 “Beginner”, 713 “Intermediate” and 512 “Expert” samples), total top-1 accuracy favors the model that performs better on the “Intermediate” or “Expert” scenario. After training on the “Expert” dataset the top-1 accuracy is significantly better on the “Expert” and the “Intermediate” test data than on the “Beginner” data. This is assumed to be due to the “Expert” training dataset being more similar to the “Expert” and “Intermediate” subsets of the test data than to the “Beginner” test data. However, the model trained on the “Expert” dataset is still able to perform reasonably well on “Beginner” data with a top-1 accuracy of 0.433. This also proves the assumption that the skeleton based approach provided a solid abstraction from the visual appearance of the workers in the original video recordings. Additional augmentation techniques may improve these accuracies further and compensate the lack of variance in the expert dataset variant.

The generally worse performance of PoseC3D on the industrial dataset could be explained by the following data annotation quality problems and the features used for the actual activity recognition. For one, some sections of the video contain ambiguous activities in respect to the defined activity classes, e.g. picking up an object with one hand and measuring a part with a tool in the other hand. This aspect can also be seen in the confusion matrix in Figure 3. It shows the misclassification behavior of the “expert” model on the industrial test dataset. Activity classes *“0 - Idle”, “1 - Walking”, “2 - Measure with tool”, “5 - Pick object”, “7 - Interact with machine interface”* and *“8 - Feed part to machine”* were mostly classified correctly. Class *“4 - Test part with hands”* was often misclassified as class *“2 - Measure with tool”*, since it was also a very stationary activity and looked similar in the footage. Class *“5 - Pick object”* and *“6 - Place object”* was often confused with each other, presumably due to it being the same activity but only in reverse. Class *“3 - Rearrange part”* was misclassified most often, which is probably due to its similarity to other classes like *5, 6* and *8*.

These results show how important choosing the correct activity classes is. Some classes were not distinguishable from each other using only skeleton data and may require image features to be distinguished. A way to incorporate image features into the pipeline was shown in [9]. In addition, a very precise definition of activities and a significant amount of samples for each of the activities is essential for good recognition results.

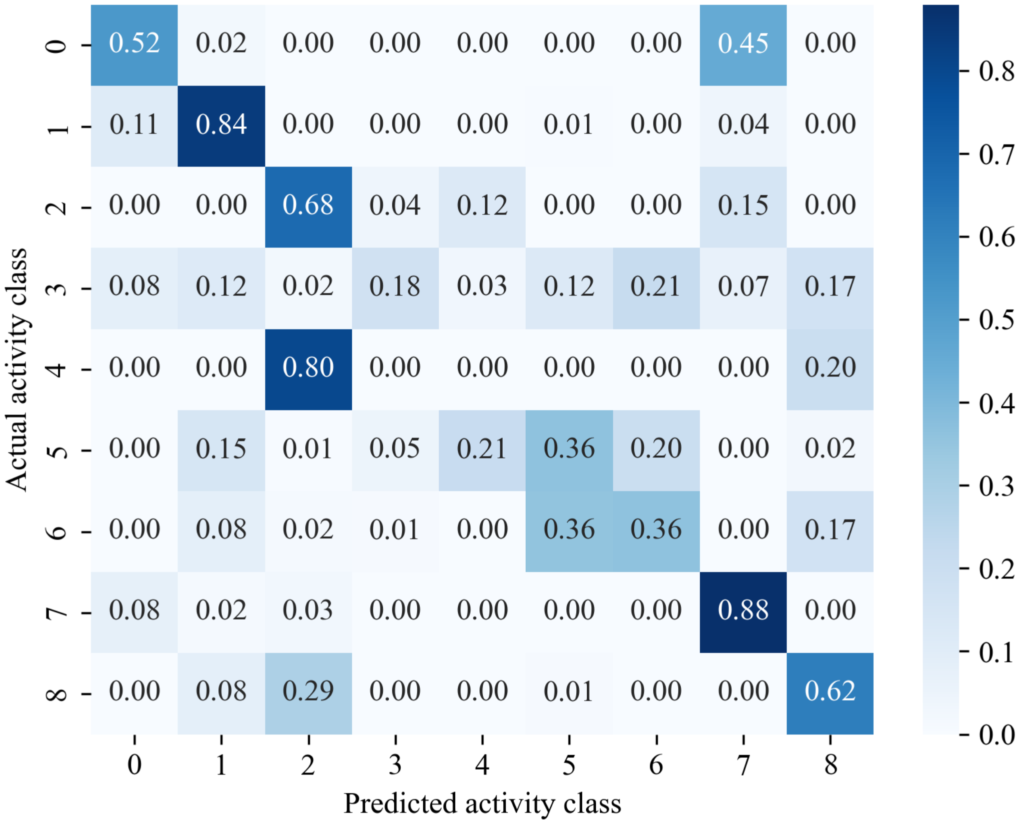


Figure 3: Confusion matrix of PoseC3D "Expert" with $I_p = 1$ and $N_{PC3D} = 48$ after 10 epochs. Cell numbers specify the ratio of occurrence in respect to the total occurrences in the specific row. Activity labels correspond to the following classes: "*0 - Idle*", "*1 - Walking*", "*2 - Measure with tool*", "*3 - Rearrange part*", "*4 - Test part with hands*", "*5 - Pick object*", "*6 - Place object*", "*7 - Interact with machine interface*", "*8 - Feed part to machine*".

4.2 PoseC3D input data variation

As PoseC3D in its baseline configuration with an input clip size of 48 is rather computationally expensive, the accuracy results and latency under multiple variations of input clip size $N_{PC3D} \in \{48, 24, 16, 8, 4\}$ and the respective pose estimation interval $I_p \in \{1, 2, 3, 6, 12\}$ were examined. These combinations were selected to retain the same time window of 1.6 s for the input clip. Depending on these parameters, the 48-frame clips were subsampled. For example, in the case of *PC3D-24-2* only every other pose frame was collected due to $I_p = 2$, resulting in 24-frame clips that correspond to the input clip size of 24. The training parameters were identical to the experiment of Section 4.1., except for the change in input clip size of the model and changing pose estimation interval in data pre-processing. The experiment was again performed on a NVIDIA Tesla V100 GPU. These model variants were trained on the "Expert" dataset, because it is more relevant to the SLEM use case. It was assumed that there is a trade-off between accuracy and latency, especially since lower latency was achieved by using fewer frames that could potentially cause the model to not receive enough data in order to make an accurate prediction.

Top-1 accuracy in the results from Table 3 look very similar across all model configurations, although *PoseC3D-16-3* achieved the top result with a top-1 total accuracy of 0.513. In contrast to that, top-5 accuracy varied more and showed *PC3D-48-1* on the high end with 0.952 and *PC3D-4-12* on the low end with 0.850. These results directly mirrored the reduction in input clip size and pose data per time interval. At last, latency results showed that the smallest model input size with $N_{PC3D} = 4$ took the least time to recognize the activities. The total latency result revealed, that only the *PC3D-4-12* variant was able to be executed in real time in the given time window of 1.6 s and on the given hardware, because the total latency did not surpass this time threshold. To lower the latency even further there are several major ways:

- Optimizing model execution using optimized inference libraries like NVIDIA's TensorRT
- Scale up hardware, e.g. with more or faster GPU (or comparable specialized hardware)
- Replace pose estimation method with a faster model combination: Faster R-CNN takes 124 ms for the detection and HRNet 93 ms for pose extraction, in contrast to 19 ms of *PC3D-4-12* for activity recognition. Replacing the R-CNN based object detector with a faster single shot detector like YOLOX [22] could improve latency significantly with a small loss in accuracy.

Table 3: Accuracy and latency results of PoseC3D with variation of input clip size N_{PC3D} and pose estimation interval I_p after 10 epochs of training on "Expert" dataset. Latency results are measured on a NVIDIA Tesla V100 GPU.

Model configuration			Top-1 accuracy	Top-5 accuracy	Latency (ms)	
Name	N_{PC3D}	I_p	Total	Total	PoseC3D	Total
PC3D-48-1	48	1	0.508	**0.952**	161	18144
PC3D-24-2	24	2	0.504	0.918	94	7464
PC3D-16-3	16	3	**0.513**	0.904	56	4368
PC3D-8-6	8	6	0.491	0.880	31	1984
PC3D-4-12	4	12	0.498	0.850	**19**	**940**

5. Conclusion

This work shows how the state-of-the-art skeleton based HAR model PoseC3D can be applied to an industrial setting to enable real time activity recognition for advanced assistance systems like SLEM. With SLEM, first steps were taken towards a system to assist new workers in machine operation and reduce teaching effort. Specifically, it was of special interest how expert knowledge can be preserved and used to guide less experienced workers. Leveraging a machine learning based approach to HAR enables the possibility to adapt to a wide variety of machine types without any reprogramming necessary for the underlying code base.

Skeleton based HAR proved useful in reducing complexity of the input data as well as abstracting away from the few and select people that are going to be observed in the industrial setting. Resulting accuracies of the conducted experiments show that even with data from only one person, skeleton based HAR is able to recognize activities of other people.

Experiments with different variations in PoseC3D's input clip sizes and the interval used for pose estimation during online activity recognition with PoseC3D make clear that inference latency can be reduced by a significant amount without noticeably affecting accuracy on the industrial dataset. It was shown which pipeline configuration is able to provide real time activity recognition.

In the future, we continue to improve on this approach by examining promising aspects. Class activities that are hard to detect with skeleton data will need further consideration, e.g. by combining high-level image features with skeleton data [9]. Additionally, we want to directly compare skeleton based HAR approaches with image based methods in the future. Optimization of the total system latency is crucial to use a HAR system in a real world application and will be further assessed. Augmentation of pose data promises to expand the available dataset and enable the model to generalize better on people with different sizes and different habits in executing activities. Finally, the amount of data required to train a robust HAR model for industrial applications must be thoroughly studied with respect to application complexity and activity classes.

Acknowledgements

The SLEM project is funded by the Baden-Württemberg Ministry of Economics, Labor and Housing as part of the Baden-Württemberg AI Innovation Competition.

References

[1] Duan, H., Zhao, Y., Chen, K., Shao, D., Lin, D., Dai, B., 2021. Revisiting Skeleton-based Action Recognition. CoRR abs/2104.13586.

[2] Ignatov, A., 2018. Real-time human activity recognition from accelerometer data using Convolutional Neural Networks. Applied Soft Computing 62, 915–922.

[3] Antar, A.D., Ahmed, M., Ahad, M.A.R., 2019. Challenges in sensor-based human activity recognition and a comparative analysis of benchmark datasets: a review, in: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), pp. 134–139.

[4] Hutchinson, M., Samsi, S., Arcand, W., Bestor, D., Bergeron, B., Byun, C., Houle, M., Hubbell, M., Jones, M., Kepner, J., Kirby, A., Michaleas, P., Milechin, L., Mullen, J., Prout, A., Rosa, A., Reuther, A., Yee, C., Gadepally, V., 2020 - 2020. Accuracy and Performance Comparison of Video Action Recognition Approaches, in: 2020 IEEE High Performance Extreme Computing Conference (HPEC). 2020 IEEE High Performance Extreme Computing Conference (HPEC), Waltham, MA, USA. 9/22/2020 - 9/24/2020. IEEE, pp. 1–8.

[5] Wang, L., Xiong, Y., Wang, Z., Qiao, Y., Lin, D., Tang, X., van Gool, L., 2019. Temporal Segment Networks for Action Recognition in Videos. IEEE Transactions on Pattern Analysis and Machine Intelligence 41 (11), 2740–2755.

[6] Feichtenhofer, C., Fan, H., Malik, J., He, K., 2019. SlowFast Networks for Video Recognition, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).

[7] Yan, S., Xiong, Y., Lin, D., 2018. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. AAAI 32 (1).

[8] Lee, J., Ahn, B., 2020. Real-Time Human Action Recognition with a Low-Cost RGB Camera and Mobile Robot Platform. Sensors 20 (10).

[9] Luvizon, D.C., Picard, D., Tabia, H., 2018. 2d/3d pose estimation and action recognition using multitask deep learning, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5137–5146.

[10]Jiao, Z., Jia, G., Cai, Y., 2020. Ensuring Computers Understand Manual Operations in Production: Deep-Learning-Based Action Recognition in Industrial Workflows. Applied Sciences 10 (3), 966.

[11]Akkaladevi, S.C., Heindl, C., 2015. Action recognition for human robot interaction in industrial applications, in: 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), pp. 94–99.

[12]Bexten, S., Schmidt, J., Walter, C., Elkmann, N., 2021. Human Action Recognition as part of a Natural Machine Operation Framework, in: 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–8.

[13]Meta AI, 2022. Papers with Code - NTU RGB+D Benchmark (Skeleton Based Action Recognition). https://paperswithcode.com/sota/skeleton-based-action-recognition-on-ntu-rgbd. Accessed 4 February 2022.

[14]MMAction2 Contributors, 2020. OpenMMLab's Next Generation Video Understanding Toolbox and Benchmark. https://github.com/open-mmlab/mmaction2. Accessed 4 February 2022.

[15]Ren, S., He, K., Girshick, R., Sun, J., 2017. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (6), 1137–1149.

[16]Sun, K., Xiao, B., Liu, D., Wang, J., 2019. Deep High-Resolution Representation Learning for Human Pose Estimation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.

[17]Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., Dai, J., Wang, J., Shi, J., Ouyang, W., Loy, C.C., Lin, D., 2019. MMDetection: Open MMLab Detection Toolbox and Benchmark. CoRR abs/1906.07155.

[18]MMPose Contributors, 2020. OpenMMLab Pose Estimation Toolbox and Benchmark. https://github.com/open-mmlab/mmpose. Accessed 4 February 2022.

[19]Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft coco: Common objects in context, in: European conference on computer vision, pp. 740–755.

[20]Wiedenroth, S.J., Denecke, J., Effenberger, I., 2020. Approach to Understand Learner Needs and Usage of the Outcome for Creativity Techniques Adressing the Learner Experience in Workplaces, in: ICERI2020 Proceedings. 13th annual International Conference of Education, Research and Innovation, Online Conference. 09.11.2020 - 11.11.2020. IATED, pp. 4365–4372.

[21]Shahroudy, A., Liu, J., Ng, T.-T., Wang, G., 2016. Ntu rgb+ d: A large scale dataset for 3d human activity analysis, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1010–1019.

[22]Ge, Z., Liu, S., Wang, F., Li, Z., Sun, J., 2021. YOLOX: Exceeding YOLO Series in 2021. CoRR abs/2107.08430.

Biography

Timo Leitritz (*1995) has been research associate at Fraunhofer Institute for Manufacturing Engineering and Automation IPA since November 2020. Timo Leitritz received a Master's degree in Mechatronics and Information technology at the Karlsruhe Institute of Technology. His main research topics are human activity recognition and AI acceleration.

Martina Köhler (*1993) has been research associate at Fraunhofer Institute for Manufacturing Engineering and Automation IPA since April 2021. Martina Köhler achieved her Masters of Science degree at the University of Ulm in Cognitive Systems. Her main research topics are human pose estimation and user experience.

Christian Jauch (*1990) has been a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart, Germany since 2015 and studied technical cybernetics at the University of Stuttgart. He works in the Image and Signal Processing department and leads the Scene Analysis group. His work focuses on hand pose estimation and activity recognition in industrial scenarios.

3rd Conference on Production Systems and Logistics

Enabling Quality-oriented Process Development for sulfidic All-Solid-State Battery Cathodes

Thorben Prein[1,2*], Hans-Christoph Töpper[1,3*], Rüdiger Daub[3]
[1]*TUMint. Energy Research GmbH, Lichtenbergstraße 4, 85748 Garching, Germany*
[2]*Technical University of Darmstadt, Karolinenplatz 5, 64289 Darmstadt, Germany*
[3]*Institute of Machine Tools and Industrial Management, Technical University of Munich, Boltzmannstraße 15, 85748 Garching, Germany*
**: Theses authors contributed equally to this work*

Abstract

After major advances in material research throughout recent years, the industrialization of all-solid-state batteries now depends on the development of cost-effective production technology for novel materials and components. To enable a fast production scale-up and complex process interdependency handling, production engineering needs a quantitative evaluation and comparison approach for manufacturing strategies and process parameter settings. To address this challenge, we derive microstructural quality criteria from specifications at the product-level such as driving range and charging speed of battery electric vehicles. These range from porosity and agglomerate density on a macroscopic level to microscopic properties such as pore size distribution and particle contacts. By listing comprehensive characterization methods, the work enables engineers to efficiently evaluate these criteria. Experimentally applying the proposed approach, the influence of different mixing process parameters is analyzed. Thereby, sulfidic composite cathodes manufactured in a scalable procedure are used as samples.

Keywords

All-solid-state battery; Battery production; Process development; Quality criteria; Electrode characterization

1. Introduction

A central issue in present-day society is the search for sustainable energy storage solutions to counteract climate change and accelerate the electrification of a wide range of products [1–4]. Due to their beneficial storage capabilities, batteries gained significant importance in the process, and forecasts predict a rapid increase in global demand in the coming years [5–7]. The strongest driver of this development is the electrification of vehicles, which requires a particularly large amount of high-quality battery cells [8, 9]. Today, the majority of these cells are conventional lithium-ion batteries (LIB) since they possess advantageous storage properties resulting from their rapid development in recent years [10–12]. Despite multiple approaches to further improve the LIB capacity by advancing the components' materials, the unsurpassable physiochemical LIB storage potential will be reached soon [12, 13]. Therefore, the only way to meet the globally growing demand is the development of novel battery generations [12, 14–19]. Among various technologies being pursued, the all-solid-state battery (ASSB) stands out in particular [12, 20–23]. ASSBs possess a dense separator made from a solid-state electrolyte featuring high mechanical stability counteracting dendrite growth at its anode interface [13, 23, 24]. This enables the use of lithium metal anodes, significantly reducing the volume and weight of the battery cell [13, 23, 25]. As a result, an increase

DOI: https://doi.org/10.15488/12119

in energy density of up to 70 % regarding the volume and up to 40 % regarding weight is predicted [12, 20, 23]. Furthermore, the solid electrolyte possesses slower calendrical aging, a higher intrinsic safety, as it cannot ignite, and in addition, is superior to the liquid electrolyte in terms of ionic conductivity [20, 23, 25, 26].

ASSB technology can be subdivided by the materials used as electrolytes [27–30]. Current research focuses on the application of polymer, oxide, and sulfide materials [31–35]. This contribution focuses on the latter, as it possesses various advantages over polymers and oxides [21, 36]. These include the electrolytes' high ionic conductivity [31, 37, 38], as well as room temperature processing in routes already established in LIB production [20, 36–38]. However, due to the low maturity of the technology and especially its currently unavailable production capabilities, sulfidic ASSBs have not been utilized in products today [39, 40]. The industrialization and application of the sulfidic all-solid-state battery now depend on the development of industrial-scale production [12, 24, 36, 38]. To support this aim regarding the sulfidic composite cathode, we derive product-specific quality criteria to systematically guide process studies toward high-quality components.

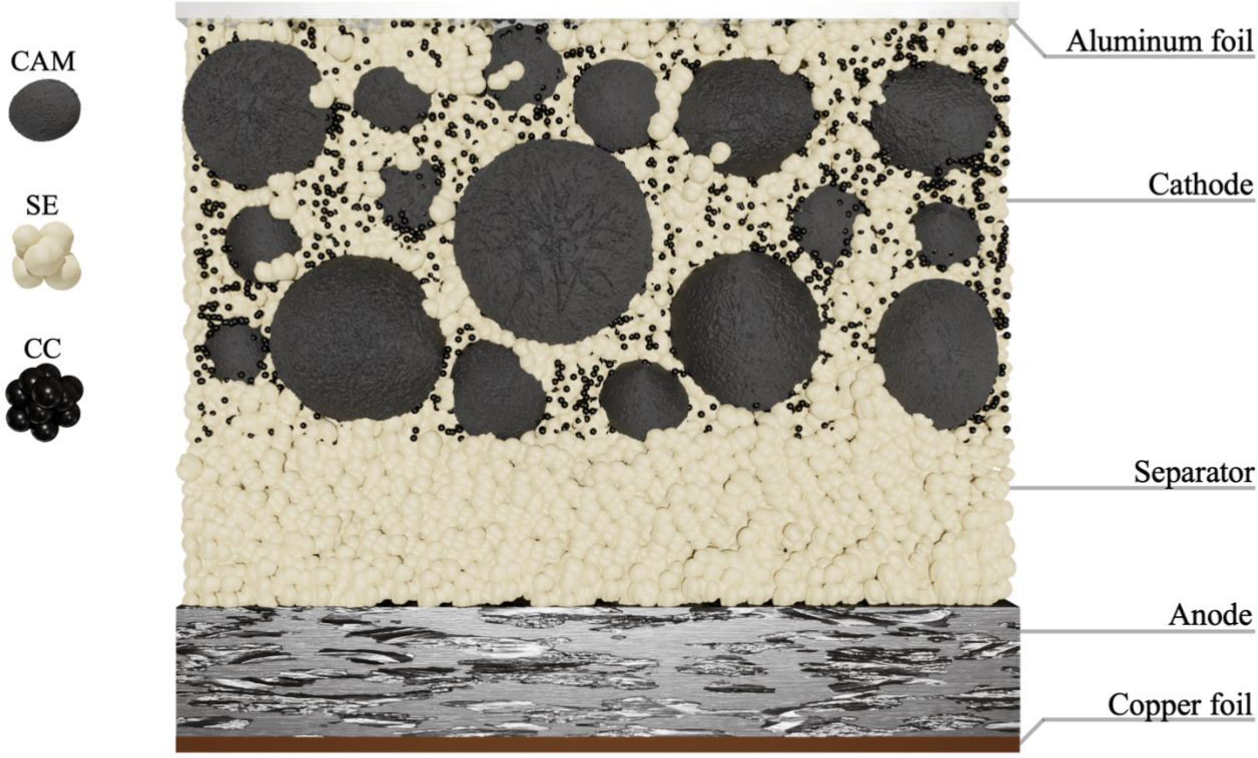


Figure 1: Schematic top-view of an ASSB galvanic cell containing cathode active material (CAM), solid electrolyte (SE), conductive carbon black (CC), SE separator and lithium metal anode.

2. Sulfidic all-solid-state battery and composite cathode

A sulfidic ASSB can be divided into multiple galvanic cells which represent the smallest electrochemically operational unit. Such a galvanic cell is depicted in figure 1. It consists of an anode, a composite cathode, and a separator in between [24, 36, 41]. While the electrodes provide the host structure for lithium-ions enabling electrical energy storage, the separator serves both as electrical insulation and an ionic connection between them. During battery charging, ions move from the composite cathode through the separator and deposit onto the anode. Simultaneously, electrons flow from the composite cathode to the anode, driven by an electrical charger outside the cell. During discharging, respective directions are reversed and a product consuming electrical energy is powered.

The sulfidic all-solid-state composite cathode is composed of an aluminum current collector for electrically contacting the electrode, lithium-ion storing cathode active material (CAM), conductive carbon (CC) improving electrical conduction within the cathode, a polymeric binder setting mechanical properties, and sulfide solid electrolyte (SE) [36, 38]. The latter is new to battery technology and specific for this type of ASSB. The SE is an ion conductor and enables lithium ions to move between the electrodes during the charging and discharging of the cell. Due to the fact, that the composite cathode is the largest and heaviest component of a galvanic cell, it possesses high influence on an ASSBs' energy density. The same correlation exists in terms of power density, as ion-conducting pathways are most critical inside the cathode. It can thus be concluded, that harnessing the full potential of sulfidic ASSB cells is only possible by managing the establishment of quality-oriented production technology for its composite cathode.

The identification of suitable manufacturing routes, stabilization of individual processes, and fine-tuning of their parameters are the focus of the composite cathode production scale-up [42–44]. According to the current state of research, slurry-based production processes already in use for manufacturing conventional electrodes may be applicable [36–38]. These include mixing, coating, drying, and calendaring. The major difference and novelty to production technique compared to LIB cathode production is the addition of the electrolyte as one of the cathodes' integral constituents at the beginning of the production process [12, 13, 45, 46]. As its production-related properties are unknown, the transfer of expertise from manufacturing LIB cathodes is limited. Accordingly, comprehensive studies are necessary for establishing suitable processes.

To provide a basic understanding of the composite cathode design choices, table 1 contains the correlation between the cathode constituents and an ASSB cells' energy and power density. The former correlates product characteristics such as the driving range of an electric vehicle, the latter influences properties such as its acceleration and charging speed.

Table 1: The influence of component-level sulfidic composite cathode design choices on product-level vehicle performance, like charging speed or acceleration and driving range is shown. Therefore, the influence of the design parameter's increase to the respective component-level quantity is indicated by ↑ (increase) or ↓ (decrease).

Criterion ↑	**Function**	**Energy density**	**Power Density**
CAM|SE ratio	Higher lithium-ion storage capacity	↑	↓
Loading	Shorter lithium-ion pathways	↓	↑
CC share	Higher electric conductivity	↓	↑
Binder share	Lower weight and volume of inactive material	↓	↓

The CAM|SE ratio describes the relative amount of cathode active material to solid electrolyte in the composite cathode. Adding more lithium-ion storing CAM and simultaneously reducing the solid electrolyte share will lead to an increase in energy density [45]. For an increase in power density, the ratio should be reduced to provide more ion-conducting electrolyte and a lower flux density in the composite cathode [22, 56]. This enables faster ion transport and therefore reduces charging and discharging times [38]. However, varying the volume fractions beyond certain thresholds may lead to either an excess of unused CAM or stagnation in power density [22, 45]. The additive shares relate to conductive carbon which provides electrical conductivity and binder enabling mechanical stability [36, 38]. Since both materials do not contribute to lithium-ion storage, their share should be as low as possible in terms of energy density increase. However, to boost power density, the proportion of CC should be increased to ensure sufficient electrical conductivity, correlating to simultaneously increasing ionic conductivity [47]. Through a cathodes' loading, the storage capacity available per unit of area is influenced. The higher the loading, the thicker the component. This leads to an increase in energy density due to the fact, that fewer separators and current collectors are incorporated given a certain cell geometry. However, since thicker electrodes result in longer ion pathways, loading should be reduced to increase power density [23, 45].

3. Microstructural quality criteria for composite cathodes

To support the development of production technology for sulfidic composite cathodes, we provided orientation for process engineers understanding of sulfidic composite cathodes regarding the influence of various constituents in the previous chapter. For a certain composite cathode composition, both energy and power density are highly influenced by production processes as they set details of the microstructure determining its performance in an ASSB [22, 45]. Consequently, the task is to optimize both criteria by finding suitable production processes and their parameters.

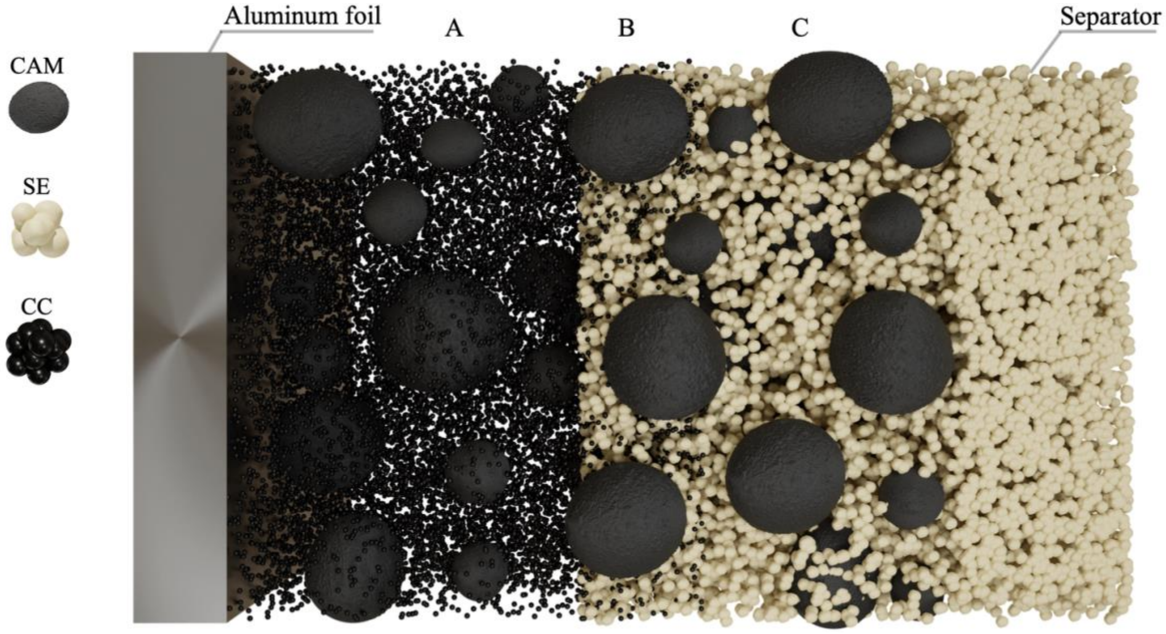


Figure 2: Schematic representation of the particle distribution within the sulfidic composite cathode. For illustration, not all constituents are shown. A) Conductive carbon black (CC) enables electron conduction between the cathode active material (CAM) particles and the aluminium conductor foil. B) CC, CAM and solid electrolyte (SE) are shown. C) SE particles enable ion transport between CAM particles and the separator.

To address this challenge new to production technology, battery cell performance determining cathode microstructure quality criteria are derived (Table 2), providing the baseline for quantitative comparison of their influence. They enable the quality-oriented identification of suitable processes and their parameters along the entire process chain. In the following, the quality criteria are listed, and explanation of their principles of action and application is given.

Porosity

A composite cathodes' porosity is defined as the pore volume in relation to the total volume [48, 49]. As described earlier in LIB technology these pores are infiltrated by the liquid electrolyte, however, for ASSB production, the electrolyte material is already added to the cathode slurry [36, 38, 50–52]. Thus, the remaining porosity represents the share of empty volume not contributing to a battery’s functions, reducing energy and power density. Therefore, porosity is associated with the lack of particle contacts, being relevant for most of the derived quality criteria. Via complex to measure but precise mercury intrusion porosimetry (MIP) spectra, the quantification of pore sizes and frequency of appearance is enabled. These can further be interlinked to the pores’ causes. In contrast, porosity-determining volume and mass measurement (VMM) is an easier way but provides no details regarding the pore size distribution. Reducing the porosity of the composite cathode to a minimum, while keeping good particle distributions is one of the main goals in composite cathode production as it significantly influences the ASSB cells' performance [45, 53].

Table 2: Sulfidic composite cathodes quality criteria applied for the evaluation of production processes and their parameters are listed. The feature optimization (Opt.) describes if a criterion is to be maximized (↑) or minimized (↓) to enhance quality. A relatively easy achievable process adaptation for a criterion's fulfillment (ful.) is indicated with +, a difficult one with -. In addition, features are classified about their cell-level functions, improving (↑), decreasing (↓), or not affecting (o) electric or ionic conductivity. The last column lists possible characterization methods of each quality criterion and gives orientation for its measurability (mea.), which is indicated by high (+) and low (-).

Quality criteria				Function		Characterization	
Material	**Criterion**	**Opt.**	**Ful.**	**Electric**	**Ionic**	**Method**	**Mea.**
-	*Porosity*	↓	-	↓	↓	VMM, MIP	+
CAM	*Homog. of distr. CAM*	↑	+	↑	↑	SEM, EDS	+
SE	*Contact SE|CAM*	↑	-	o	↑	SEM, MIP	+
	Contact SE|SE					SEM, MIP	+
	Homog. of distr. SE					SEM, EDS, MIP	+
CC	*Contact CC|CAM*	↑	+	↑	o	SEM	-
	Contact CC|CC					SEM, MIP	-
	Homog. of distr. CC					SEM, EDS, MIP	+

Cathode active material

The CAM particles are starting point of electron and ion pathways through a cathode's microstructure [56]. That is why their distribution and connection are influential to the electrical and ionic conductivity of composite cathodes. The homogenous distribution of CAM enables SE and CC particles to fill gaps between the larger CAM particles. This counteracts the formation of CAM particles separated from the conduction cluster and therefore ensures high CAM utilization rates. Furthermore, the homogeneous distribution, lowers both, ion and electron pathway length and flux density, ultimately decreasing resistances [45, 53-55]. Process wise the distribution of CAM particles is rather easy to achieve but is linked to SE particle distribution and can therefore be aggravated. To determine the criterion, scanning electron microscopy (SEM) can be used. CAM particles are easily identified on top-view images of cross-sections, the estimation of their distribution can be conducted qualitatively by humans. If needed Energy-dispersive X-ray spectroscopy (EDS) further simplifies CAM particle identification. Because of its central functional role, the homogeneity of CAM particle distribution is an important quality criterion.

Solid Electrolyte

During charging and discharging, solid electrolyte particles are constituting the conductive element for ion transport. The interrelation between a composite cathode's microstructural geometry on the ionic conductivity is characterized by the tortuosity factor [57]. This describes the ratio of the length of the effective lithium-ion pathway to the shortest possible pathway. It should be as low as possible for an even flux density and low ionic resistance. To increase ion conduction and therefore keep the tortuosity factor low, the active interface area formed by CAM and SE particles contacts are crucial as they allow lithium ions to enter and leave CAM particles during charging and discharging. The more pronounced the contact, the lower the ionic interface resistance [45, 53-56]. Process-wise, this parameter is very difficult to adjust. Due to differences in the size of the SE and CAM particles, contacts can be approximated by MIP and identified using SEM. In addition to SE and CAM particle contact, individual SE particles must be in good contact to reduce the ionic resistance of the composite cathode [40]. Due to the easily detectable SE particles, the fulfillment of this criterion can be evaluated using SEM images or detected via MIP spectra. Furthermore, unfavorable SE particle distribution leads to increased tortuosity, and thus lowers the composite cathode ionic conductivity. This results in lower power densities through slower ion transfer and lower energy densities because of inactive CAM particle fractions. It can be qualitatively analyzed rather precisely by humans via particle detection on top-view SEM images as well as cross-sections.

Conductive carbon

To increase the composite cathodes' electrical conductivity resulting from CAM particles, graphite-based CC is added to conduct electrons between the CAM particles and aluminum conductor foil [47]. To achieve the lowest possible electrical resistance, the contact between CAM and CC, as well as between different CC particles should be as pronounced as possible. Similar to the distribution of CAM and SE, the homogenous CC particle arrangement ensures sufficient electrical conduction, lowering electrical resistance [57]. Process-wise, the relatively small carbon particles are easily distributable, but can hardly be identified by SEM images or mercury intrusion porosimetry spectra. Accordingly, EDS analysis is used to identify carbon atoms' locations.

4. Procedure for applying quality criteria to production processes

Since knowledge of the production processes is currently unavailable and expertise from the manufacturing of conventional batteries is hardly transferable, an iterative approach must be taken to design suitable processing routes. Thereby, the knowledge of the quality criteria and their characterization plays a central role. Because only through the systematic comparison of their fulfillment, the influence of routes, processes, and parameters becomes quantitatively comparable.

When applying the quality criteria for production development, a total of three stages must be passed.

1. At first, the object of investigation must be chosen. This may be an entire manufacturing route consisting of multiple processes, an individual step composed of various parameters, or the variation of just a single process parameter.
2. Composite cathodes must be produced, and the quality criteria to be investigated chosen. The focus can be on all criteria at once, but of course also refer to only one single aspect, which is determined by the stage of process development. The former tends to be suited for entire manufacturing routes and initial production trials, the latter for the fine-tuning of individual steps and their parameters. After cathode production, the respective characterization methods are applied.
3. In the last step, the evaluation regarding the fulfillment of quality criteria is conducted. The application of this approach is independent of the production scale under consideration. Thus, it suits both the manual manufacturing of both laboratory-scale components as well as the production of large-format composite cathodes in industrially relevant processes.

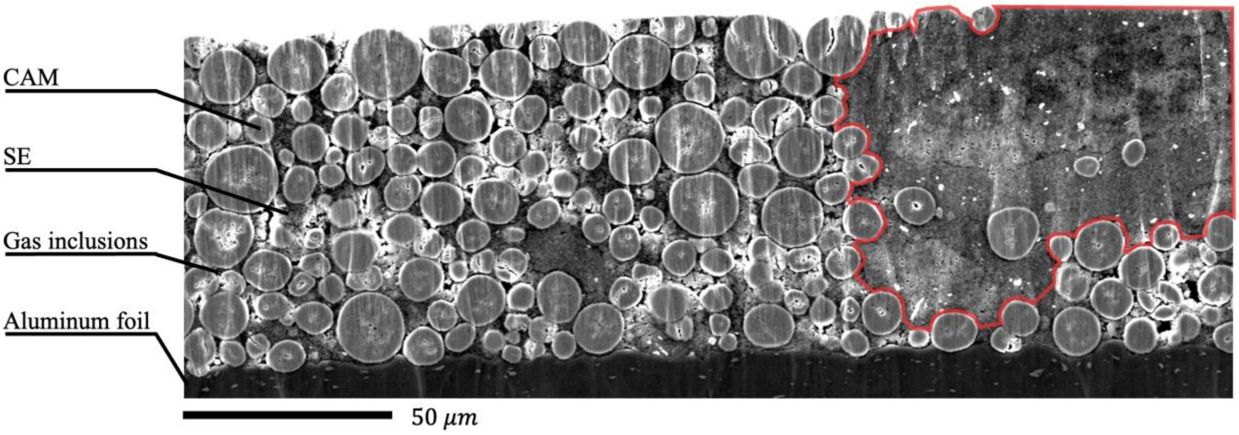


Figure 3: Cross-section of a densified sulfidic composite cathode after processing in a wet chemical route at the Institute of Machine Tools and Industrial Management. On the top right, an SE agglomerate is highlighted in red. The composite cathodes remaining porosity originates from voids between particles.

5. Exemplary analysis of quality criteria for process development

To demonstrate our approach, it is applied for the analysis of a scalable wet chemical process chain. The chain is built from mixing, coating, drying, and pressing steps which are executed in a similar fashion already described by Sakuda et al. [38]. Also, the composite cathodes material composition is adopted.

The aim of the exemplary analysis is the identification of the influence of the mixing duration on the quality criteria porosity and particle distribution before and after the pressing step. Therefore, the mixing duration is set to 2, 32, and 64 minutes, while all other process parameters as well as the material composition are kept constant. After producing uncompressed uand compressed composite cathodes with varying mixing durations, VMM, SEM, and EDS analyses are carried out to analyze the quality criteria.

For the mixing time of two minutes, the resulting slurry possesses large agglomerates. Therefore, it is not further processed. Measuring the porosity using VMM measurements it is indicated that for mixing times of 64 minutes the resulting cathodes porosity before and after compaction is higher compared to 32 minutes mixed composite cathodes. The SEM images show potential explanation in higher SE agglomeration for the longer mixing times. For all process parameter sets, agglomerates are visible. Through the SE agglomerates the CAM and SE distributions are negatively affected, as visible in figure 4 column A. The local SE exaggeration leads to the formation of SE poorer regions, possessing a high CAM share and mediocre CAM|SE contacts. Tough through compaction CAM|SE and SE|SE contacts are improved by SE deformation; distributions are not. SEM images indicate that for the 2 min mixed processing route much worse SE|SE contacts are achieved compared to the other two parameter sets. The CC distribution and contacts can be investigated by the identification of CC particles on SEM images in combination with EDS analysis. The results show a good distribution for CC particles for all three investigated parameter sets, highlighting the rather easy manageability of its distribution during composite cathode slurry mixing. Conclusively the results indicate poor fulfillment of the derived quality criteria for all three cathodes. This is indicated by high porosities and low distribution and contact qualities for CAM and SE, which are expected to lead to reduced ionic conductivities. Though the cause of agglomerate occurrence cannot be deduced, quality criteria can be used for further hypothesis-driven experimentation and evaluation. These could potentially investigate if the agglomerates are induced by the supplied SE materials and give further insights into whether they grow through the mixing process.

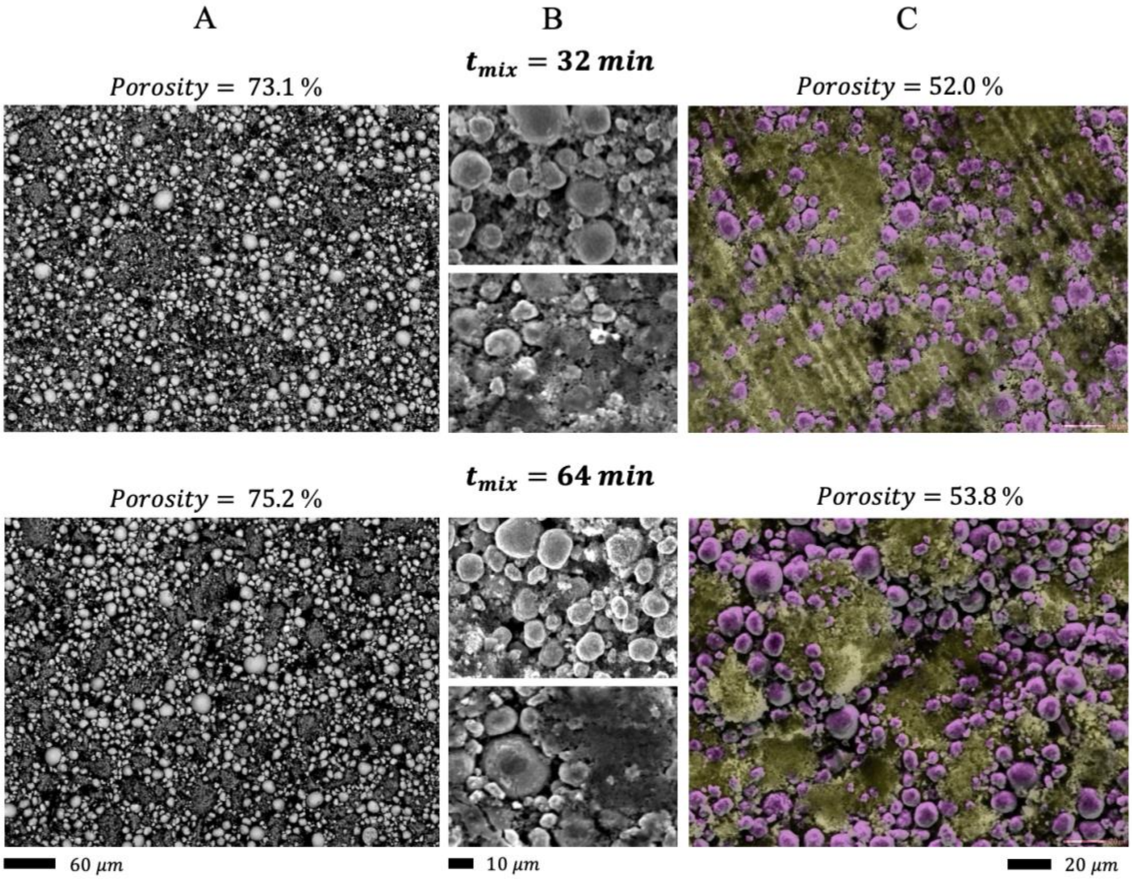


Figure 4: In column A backscatter SEM images of the uncompacted composite cathodes are shown, especially highlighting CAM distribution. In column B, secondary electron detector images indicate the particle contacts, associated with the non-compressed and compressed cathodes. The third column shows EDS analysis indicating the CAM and SE distribution after the compaction step, the active material is marked purple, while SE olive. Porosities are calculated using averaged VMM measurements for 8 samples, scales are introduced per column.

6. Conclusion and outlook

Sulfidic ASSBs represent a promising candidate for the next generation of electrochemical energy storage systems. However, industrial-scale production technology necessary for their application is currently unavailable. To support the identification of suitable manufacturing routes, processes, and their parameters, microstructural quality criteria are presented focusing on composite ASSB cathodes. For each criterion, possible measurement techniques are listed for their quantification. An exemplary use case focusing on the evaluation of the mixing durations influence on selected criteria is described to show the approach's applicability.

Further investigations may focus on improving and automating quantification methods for the proposed quality criteria. In the long term, this could be used as feedback in an automated process parameter optimization loop. In addition, aspects such as the sensitivity of the quality criteria regarding process steps and their parameters may be analyzed.

Acknowledgment

The authors express their sincere gratitude to the Bavarian Ministry of Economic Affairs, Regional Development and Energy for the funding of the research project "Industrialisierbarkeit von Festkörperelektrolytzellen".

References

[1] Gulagi, A., Bogdanov, D., Breyer, C., 2018. The role of storage technologies in energy transition pathways towards achieving a fully sustainable energy system for India 17, p. 525.

[2] Schainker, R.B., 2004. Executive overview: energy storage options for a sustainable energy future, in IEEE Power Engineering Society General Meeting, 2004, IEEE, p. 2310.

[3] Rothgang, S., Rogge, M., Becker, J., Sauer, D., 2015. Battery Design for Successful Electrification in Public Transport 8, p. 6715.

[4] Väyrynen, A., Salminen, J., 2012. Lithium ion battery production 46, p. 80.

[5] Choi, S., Wang, G., 2018. Advanced Lithium-Ion Batteries for Practical Applications: Technology, Development, and Future Perspectives 3, p. 1700376.

[6] Vaalma, C., Buchholz, D., Weil, M., Passerini, S., 2018. A cost and resource analysis of sodium-ion batteries 3, p. 652.

[7] Zhang, J., Wang, Z., Liu, P., Zhang, Z., 2020. Energy consumption analysis and prediction of electric vehicles based on real-world driving data 275, p. 115408.

[8] Jensen, A.F., Cherchi, E., Mabit, S.L., Ortúzar, J.d.D., 2017. Predicting the Potential Market for Electric Vehicles 51, p. 427.

[9] Xu, C., Dai, Q., Gaines, L., Hu, M. et al., 2020. Future material demand for automotive lithium-based batteries 1, p. 437.

[10] Kwade, A., Haselrieder, W., Leithoff, R., Modlinger, A. et al., 2018. Current status and challenges for automotive battery production technologies 3, p. 290.

[11] Jaguemont, J., Boulon, L., Dubé, Y., 2016. A comprehensive review of lithium-ion batteries used in hybrid and electric vehicles at cold temperatures 164, p. 99.

[12] Janek, J., Zeier, W.G., 2016. A solid future for battery development 1, p. 1167.

[13] Varzi, A., Raccichini, R., Passerini, S., Scrosati, B., 2016. Challenges and prospects of the role of solid electrolytes in the revitalization of lithium metal batteries 4, p. 17251.

[14] Kim, T., Song, W., Son, D.-Y., Ono, L.K. et al., 2019. Lithium-ion batteries: outlook on present, future, and hybridized technologies 7, p. 2942.

[15] Girishkumar, G., McCloskey, B., Luntz, A.C., Swanson, S. et al., 2010. Lithium−Air Battery: Promise and Challenges 1, p. 2193.

[16] Zhang, W., Zhang, F., Ming, F., Alshareef, H.N., 2019. Sodium-ion battery anodes: Status and future trends 1, p. 100012.

[17] Liu, Q., Hu, Z., Li, W., Zou, C. et al., 2021. Sodium transition metal oxides: the preferred cathode choice for future sodium-ion batteries? 14, p. 158.

[18] Hwang, J.-Y., Myung, S.-T., Sun, Y.-K., 2017. Sodium-ion batteries: present and future. Chem Soc Rev 46, p. 3529.

[19] Geng, D., Ding, N., Hor, T.S.A., Chien, S.W. et al., 2016. From Lithium-Oxygen to Lithium-Air Batteries: Challenges and Opportunities 6, p. 1502164.

[20] Wu, F., Fitzhugh, W., Ye, L., Ning, J. et al., 2018. Advanced sulfide solid electrolyte by core-shell structural design. Nat Commun 9, p. 4037.

[21] Krauskopf, T., Richter, F.H., Zeier, W.G., Janek, J., 2020. Physicochemical Concepts of the Lithium Metal Anode in Solid-State Batteries. Chem Rev 120, p. 7745.

[22] Kato, Y., Hori, S., Saito, T., Suzuki, K. et al., 2016. High-power all-solid-state batteries using sulfide superionic conductors 1, p. 652.

[23] Luntz, A.C., Voss, J., Reuter, K., 2015. Interfacial challenges in solid-state Li ion batteries 6, p. 4599.

[24] Kim, M.-J., Park, J.-W., Kim, B.G., Lee, Y.-J. et al., 2020. Facile fabrication of solution-processed solid-electrolytes for high-energy-density all-solid-state-batteries by enhanced interfacial contact. Sci Rep 10, p. 11923.

[25] Froboese, L., Groffmann, L., Monsees, F., Helmers, L. et al., 2020. Enhancing the Lithium Ion Conductivity of an All Solid-State Electrolyte via Dry and Solvent-Free Scalable Series Production Processes 167, p. 20558.

[26] Hayashi, A., Masuzawa, N., Yubuchi, S., Tsuji, F. et al., 2019. A sodium-ion sulfide solid electrolyte with unprecedented conductivity at room temperature. Nat Commun 10, p. 5266.

[27] Ito, S., Fujiki, S., Yamada, T., Aihara, Y. et al., 2014. A rocking chair type all-solid-state lithium ion battery adopting Li2O–ZrO2 coated LiNi0.8Co0.15Al0.05O2 and a sulfide based electrolyte 248, p. 943.

[28] Kotobuki, M., Munakata, H., Kanamura, K., Sato, Y. et al., 2010. Compatibility of Li[sub 7]La[sub 3]Zr[sub 2]O[sub 12] Solid Electrolyte to All-Solid-State Battery Using Li Metal Anode 157, A1076.

[29] Agrawal, R.C., Hashmi, S.A., Pandey, G.P., 2007. Electrochemical cell performance studies on all-solid-state battery using nano-composite polymer electrolyte membrane 13, p. 295.

[30] Jiang, Y., Yan, X., Ma, Z., Mei, P. et al., 2018. Development of the PEO Based Solid Polymer Electrolytes for All-Solid State Lithium Ion Batteries. Polymers (Basel) 10.

[31] Ohtomo, T., Hayashi, A., Tatsumisago, M., Tsuchida, Y. et al., 2013. All-solid-state lithium secondary batteries using the 75Li2S·25P2S5 glass and the 70Li2S·30P2S5 glass–ceramic as solid electrolytes 233, p. 231.

[32] Yu, Q., Han, D., Lu, Q., He, Y.-B. et al., 2019. Constructing Effective Interfaces for Li1.5Al0.5Ge1.5(PO4)3 Pellets To Achieve Room-Temperature Hybrid Solid-State Lithium Metal Batteries. ACS Appl Mater Interfaces 11, p. 9911.

[33] Angulakshmi, N., Nahm, K.S., Nair, J.R., Gerbaldi, C. et al., 2013. Cycling profile of MgAl2O4-incorporated composite electrolytes composed of PEO and LiPF6 for lithium polymer batteries 90, p. 179.

[34] Han, P., Zhu, Y., Liu, J., 2015. An all-solid-state lithium ion battery electrolyte membrane fabricated by hot-pressing method 284, p. 459.

[35] Kimura, K., Yajima, M., Tominaga, Y., 2016. A highly-concentrated poly(ethylene carbonate)-based electrolyte for all-solid-state Li battery working at room temperature 66, p. 46.

[36] Ates, T., Keller, M., Kulisch, J., Adermann, T. et al., 2019. Development of an all-solid-state lithium battery by slurry-coating procedures using a sulfidic electrolyte 17, p. 204.

[37] Lee, K., Kim, S., Park, J., Park, S.H. et al., 2017. Selection of Binder and Solvent for Solution-Processed All-Solid-State Battery 164, A2075-A2081.

[38] Sakuda, A., Kuratani, K., Yamamoto, M., Takahashi, M. et al., 2017. All-Solid-State Battery Electrode Sheets Prepared by a Slurry Coating Process 164, A2474-A2478.

[39] Nam, Y.J., Cho, S.-J., Oh, D.Y., Lim, J.-M. et al., 2015. Bendable and thin sulfide solid electrolyte film: a new electrolyte opportunity for free-standing and stackable high-energy all-solid-state lithium-ion batteries. Nano Lett 15, p. 3317.

[40] Takahashi, K., Hattori, K., Yamazaki, T., Takada, K. et al., 2013. All-solid-state lithium battery with LiBH4 solid electrolyte 226, p. 61.

[41] Sakuda, A., Hayashi, A., Ohtomo, T., Hama, S. et al., 2011. All-solid-state lithium secondary batteries using LiCoO2 particles with pulsed laser deposition coatings of Li2S–P2S5 solid electrolytes 196, p. 6735.

[42] Tan, D.H.S., Banerjee, A., Chen, Z., Meng, Y.S., 2020. From nanoscale interface characterization to sustainable energy storage using all-solid-state batteries. Nat Nanotechnol 15, p. 170.

[43] Schnell, J., Knörzer, H., Imbsweiler, A.J., Reinhart, G., 2020. Solid versus Liquid—A Bottom-Up Calculation Model to Analyze the Manufacturing Cost of Future High-Energy Batteries 8, p. 1901237.

[44] Duffner, F., Kronemeyer, N., Tübke, J., Leker, J. et al., 2021. Post-lithium-ion battery cell production and its compatibility with lithium-ion cell production infrastructure 6, p. 123.

[45] Bielefeld, A., Weber, D.A., Janek, J., 2019. Microstructural Modeling of Composite Cathodes for All-Solid-State Batteries 123, p. 1626.

[46] Lenze, G., Bockholt, H., Schilcher, C., Froböse, L. et al., 2018. Impacts of Variations in Manufacturing Parameters on Performance of Lithium-Ion-Batteries 165, A314-A322.

[47] Deng, S., Sun, Y., Li, X., Ren, Z. et al., 2020. Eliminating the Detrimental Effects of Conductive Agents in Sulfide-Based Solid-State Batteries 5, p. 1243.

[48] Günther, T., Schreiner, D., Metkar, A., Meyer, C. et al., 2020. Classification of Calendering-Induced Electrode Defects and Their Influence on Subsequent Processes of Lithium-Ion Battery Production 8, p. 1900026.

[49] Meyer, C., Kosfeld, M., Haselrieder, W., Kwade, A., 2018. Process modeling of the electrode calendering of lithium-ion batteries regarding variation of cathode active materials and mass loadings 18, p. 371.

[50] Wenzel, V., Nirschl, H., Nötzel, D., 2015. Challenges in Lithium-Ion-Battery Slurry Preparation and Potential of Modifying Electrode Structures by Different Mixing Processes 3, p. 692.

[51] Bitsch, B., Dittmann, J., Schmitt, M., Scharfer, P. et al., 2014. A novel slurry concept for the fabrication of lithium-ion battery electrodes with beneficial properties 265, p. 81.

[52] Westphal, B.G., Kwade, A., 2018. Critical electrode properties and drying conditions causing component segregation in graphitic anodes for lithium-ion batteries 18, p. 509.

[53] Hlushkou, Dzmitry, et al., 2018. The influence of void space on ion transport in a composite cathode for all-solid-state batteries 396, p. 363.

[54] Siroma, Zyun, et al. 2016. AC impedance analysis of ionic and electronic conductivities in electrode mixture layers for an all-solid-state lithium-ion battery 316, p. 215.

[55] Koerver, Raimund, et al. 2017. Capacity fade in solid-state batteries: interphase formation and chemomechanical processes in nickel-rich layered oxide cathodes and lithiumthiophosphate solid electrolytes 29.13, p. 5574.

[56] Zhang, Wenbo, et al. 2017. Interfacial processes and influence of composite cathode microstructure controlling the performance of all-solid-state lithium batteries. ACS applied materials & interfaces 9.21, p. 17835.

[57] Randau, Simon, et al. 2021. On the Additive Microstructure in Composite Cathodes and Alumina-Coated Carbon Microwires for Improved All-Solid-State Batteries 33.4, p. 1380.

Biography

Thorben Prein is a graduate student in mechanical engineering at the Technical University of Darmstadt and is currently working as student research assistant at TUMint. Energy Research GmbH and Institute for Machine Tools and Industrial Management (*iwb)* at the Technical University of Munich.

Hans-Christoph Töpper graduated in mechanical engineering at the Karlsruhe Institute of Technology and is now working as a research associate at the TUMint. Energy Research GmbH and Institute for Machine Tools and Industrial Management (*iwb)* at the Technical University of Munich.

Rüdiger Daub has been head of the Chair of Production Technology and Energy Storage Systems at *iwb* at Technical University of Munich and at the Fraunhofer Institute for Foundry, Composites and Processing Technology IGCV since 2021. Previously Prof. Dr.-Ing. Rüdiger Daub has held several positions at BMW Group Munich, most recently as head of technology development and prototype construction for lithium-ion cells.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2022

3rd Conference on Production Systems and Logistics

Modularized Active Learning Solution For Labelling Text Data For Business Environment Analysis

Furkan Agacayaklar[1], Annika Lange[1], Julia-Anne Scholz[1], Thomas Knothe[1], Dirk Busse[2]
[1]Fraunhofer Institute for Production Systems and Design Technology IPK, Berlin, Germany
[2]budatec GmbH, Berlin, Germany

Abstract

In today's interconnected world, the pace of change is increasing gradually and the effects of an event can propagate and disrupt industries, organizations or companies more dramatically and quickly. Therefore, having a comprehensive overview of the environment is a precious asset for resilience and sustainable growth. One enabler of the above-mentioned interconnectedness is the rapid flow and vast availability of information in text form, which can be also used as the fundamental resource to understand the shifting environment. Hence, actors can be able to become aware of changes at an early stage. The underlying patterns to filter relevant information can be detected by learning from data, or more specifically machine learning. Natural language processing (NLP) techniques can be applied because text data is analyzed. However, to embed the expertise and perspective of the user into the initial model, data should be labeled. This requires valuable expert time from the organization for the labeling, thus it should be minimized. This study aims to present an efficient and user-friendly solution for data labeling. To achieve this, a modularized Active Learning-based backend is combined with an intuitive interface. The output of this labeling process will be used further to train a model for environment analysis. Nevertheless, the main focus of this paper is the development of a solution to maximize efficiency during data labeling for environment analysis. After an introduction to the problem, the overview of the suggested solution accompanied by a prototype will be demonstrated.

Keywords

Business Environment Analysis; Active Learning; Natural Language Processing; Machine Learning; Data Labeling

1. Introduction

A critical challenge for companies is planning future-oriented strategies, especially since today's environment is characterized by transformations [1]. Unpredictable, disruptive events (e.g. the Corona crisis) and dynamic developments are becoming increasingly difficult for companies to assess. Moreover, disruptive influences often put companies in the situation of having to make decisions with long-term effects within a short period of time [2]. Trade barriers, for example, have a direct influence on the supply chain and can force a company to quickly decide on alternative ways to source resources under uncertainty.

Foresight and corporate environment analysis are key to identifying risks and opportunities and ensuring the long-term competitiveness of the company [3]. JAIN [4] identified the connection of environmental influences with the business strategy as an essential component for a future-oriented, agile company. To this end, making trend predictions and differentiating between relevant and irrelevant environmental influences

DOI: https://doi.org/10.15488/12190

are elementary components [4]. However, especially in small and medium-sized enterprises (SMEs), there are various barriers to carrying out environmental analysis, including high personnel and financial costs [5,6]. Moreover, the complexity of dynamic influences requires a solid overview of relevant information for short- and long-term decision-making. For this reason, the Fraunhofer Institute for Production Systems and Design Technology (IPK) has developed a model-based interactive situation awareness monitor and liquidity assessment for enterprise resilience management to provide SMEs in particular with a well-founded and company-specific basis for decision-making, even in times of crisis. In addition to information on the corporate environment, information on suppliers, their production and orders are visualized. The aim of this study is to provide content for a news ticker which is displayed on the situation awareness monitor. [5]

Consideration of context while displaying relevant information from the corporate environment requires analyzing data intelligently instead of providing unfiltered data from certain resources. This could be realized by Machine Learning (ML), more specifically Natural Language Processing [7], if data is in text format. However, this context filter model should be tailored to each organization. The perspective of the organization can be transferred into the ML model by labeling data. One or more individuals from the organization should label a various number of information based on relevancy. However, this task is manual and time-consuming, which is a problem especially when expert time is involved. To minimize the effect of this barrier, an Active Learning (AL) based labeling solution is developed which will be presented in this paper. This labeled data can be used to develop the initial model, which can be utilized for further implementation in the environment analysis pipeline. Nevertheless, the scope of this paper is concentrated on data labeling and Active Learning. Additional steps in the environment analysis pipeline such as training and performance of NLP models are part of further research and therefore not covered in this paper.

The logic behind the approach using text data, supervised ML and Active Learning for environment analysis is addressed in section 2. Next, existing approaches from Active Learning are presented (section 3), followed by the application (section 4) and discussion (section 5). The paper concludes with a summary and outlook (section 6).

2. Problem Description

2.1 Input Data for Environment Analysis

Data can take different forms. A popular depiction of data is in tabular form, which is known as structured data. These types of data can be sorted, searched, modified and analyzed via conventional methods. However, they constitute a small portion of data because most of the data created today are unstructured in the form of text, image, sound, etc. [8]. This also applies to environment analysis, since real-time information coming from the outside world (e.g. news articles) is generally in raw format and not always pre-processed or put into the structured format. Therefore, analysis of unstructured data is key for environment analysis.

Text is an informative and effective form of unstructured data, considering news, tweets and other written media. This is how people consume information. Other forms such as images or videos are also valuable sources, but even those are captioned by text. For example, a picture of a ship lodged on the ground may not have much meaning. However, if this picture is captioned as *"The giant ship is causing a traffic jam in one of the world's busiest waterways"* [9] in coverage of the famous Suez Canal blockage, it has much more context and meaning to a decision-maker. Moreover, text data is relatively lightweight compared to images and video. Therefore, text data from news providers will be considered in this study.

2.2 Requirement for Machine Learning

The relevance of a text can be intuitively evaluated by a human being, but it is not straightforward for a machine to mimic this behavior. The first option to consider is rule-based methods such as regex-based filter or ElasticSearch [10]. The required information can be found by using keywords or search queries. Similarly, certain rules can be established to filter relevant information. However, some problems may arise. First of all, it is a static solution and does not learn from experience. Secondly, some information or themes can be overlooked because in a basic search it is known what to look for, but developing a filter to find relevant information is a broader task. This is where ML comes into consideration. Using ML allows one to inherit complex patterns in the data and develop more capable solutions. Therefore, an ML model will be trained to filter relevant news. The model can be considered as the context filter which is the first step of the pipeline and chained with further models or even rule-based filters for further refinement.

The problem could be described as a text classification problem but it can be approached as both supervised or unsupervised learning. Supervised learning methods learn from labeled data (i.e. input-output pairs) so that new inputs can be mapped to respective outputs [11]. In the case of environment analysis, tailored solutions should be developed for each organization. Thus, to apply supervised learning, data should be labeled based on the preferences of the corresponding organization. This prerequisite for supervised learning is an incentive to consider unsupervised learning.

Unsupervised learning algorithms try to discover patterns in data without labels [11]. For this problem, clustering algorithms [12], which is a sub-field of unsupervised learning, could be used to divide data into groups based on similarities of data. Each cluster can be analyzed and labeled based on relevancy and new data can be evaluated based on the cluster it is mapped to. However, dividing data into ideal, desired and homogeneous clusters is a challenge and tuning of a clustering algorithm requires effort which can be illustrated by varying results of different clustering algorithms on the same dataset [12]. Therefore, supervised learning is the more attractive option despite the requirement of labeling.

2.3 Data Labeling

In the case of environment analysis, one or more people from an organization with sufficient and comprehensive domain knowledge should evaluate the selected text data based on relevancy and label the data. This means a considerable time and effort of experienced employees should be dedicated to such a manual task. Therefore, an effective and efficient choice of data to be labeled is vital. Considering the variety of news providers and the amount of news published daily, the raw pool can be large. Titles and short descriptions of 29,040 articles from various German news providers are acquired as text data for environment analysis. The next step is to select the data to be labeled.

The first and the simplest option is random sampling. The problem with this approach is that it can lead to an unbalanced data set, which means the data could include many articles which will be labeled as irrelevant. This is detrimental for ML model training [13]. However, this phenomenon is highly likely for environment analysis because the articles which are specifically relevant for an organization will be small among all themes in the article pool such as sport, entertainment and technology. Assuming only 1,000 out of 29,040 articles were relevant, on average 35 articles are expected to be relevant in a sample of 1,000. The second option is inspired by how people look up information on the Web in daily life. The information is filtered by search queries or rule-based filters. Then, the desired information is searched within this distilled list of options. Similarly, the articles can be filtered based on one or more words and filtered articles could be labeled. However, that might result in overlooking some articles or themes as mentioned before. Another option is Active Learning [14–16] which is a method used to address the labeling problem in ML and provides informed choices on data for labeling and is elaborated in the following section.

3. State of the Art of Active Learning

Active Learning aims to make informed queries from unlabeled data based on the output of learning algorithms so that labeling of these queries will allow the model to perform better [14,16]. Therefore, target performance can be reached with fewer queries and faster compared to random selection [15,16]. This aspect is valuable for environment analysis data labeling since valuable expert time is minimized by fewer data to label.

The general procedure of AL is provided in Figure 1. It starts with small set of already labeled data which is enough to train a learning algorithm. These initial data can be selected by diversity sampling strategies (will be covered) or randomly and be labeled. This is depicted by the dashed line in the figure. This labeled (initial) data is used to train the model/learning algorithm. Then, the trained algorithm selects (queries) data from the unlabeled (pool) dataset, labeling of which is considered valuable by the algorithm for the goal being sought. The term "*goal*" is mentioned here because it is generally stated as accuracy in literature [16]. However, it can differ based on the type of query strategy, which will be covered. Selected data for labeling is examined by the user or expert, which is named as "*oracle*" in AL literature [14,16]. The oracle labels the data which is transferred from the pool dataset to the initial dataset, which is generally done automatically. Afterwards, the algorithm can be refitted based on the updated version of the labeled dataset, which allows it to make more informed choices. The updated algorithm can query the new data and the cycle repeats itself. In the environment analysis example, the algorithm selects an article from the unlabeled news dataset which is shown to the user. The user only reads the article and labels it as relevant(yes) or irrelevant(no). The labeled article can be removed from the pool dataset and appended to the labeled dataset in the background automatically. The algorithm can be retrained with labeled dataset and can query the next article. Although data can have more than two dimensions, the plot on the right shows an example in 2D just for visualization. The queried data point and labeled data can be observed.

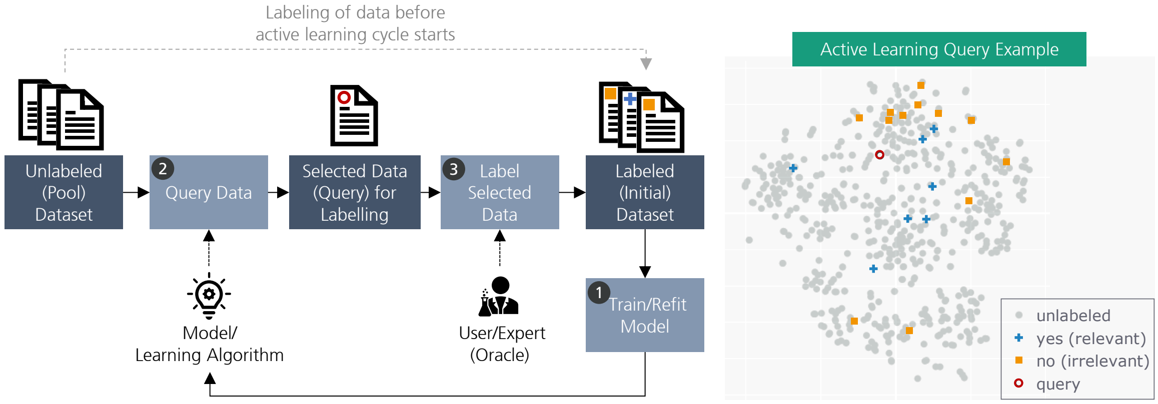


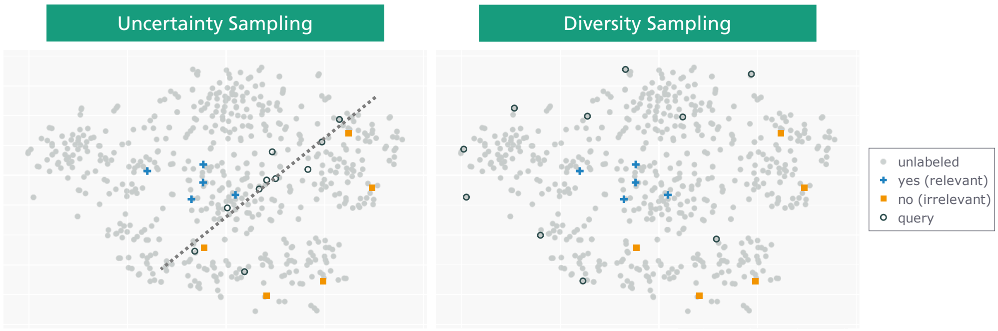
Figure 1: Generic flow of Active Learning and its representation on 2D scatter plot. Data were taken from the BBC News dataset [17]. Vectorized articles are projected on 2D via t-SNE [18] just for visualization purposes.

The goal, which is tried to be achieved by each query, can differ as mentioned before and AL query strategies can be classified based on this aspect. Settles [16] provides a good survey on query strategies, but they are not grouped. However, according to Kumar, Gupta [14], query strategies can be classified as informative-based and representative-based for classification problems. Munro [15] calls them uncertainty sampling and diversity sampling respectively, although uncertainty sampling is considered as a sub-category of informative-based strategies [14]. Nevertheless, informative-based strategies also focus on uncertainty, so underlying aims and logics are similar [14]. In this article, uncertainty sampling and diversity sampling are used for classification.

The knowledge quadrant by Munro [15] can be used to clarify the goals of strategies. Every data point belongs to one of the quadrants at a state which can be explained in the context of environment analysis. The first quadrant focuses on "*Known-knowns*", i.e. ML model knows that it can predict those data points (e.g. news articles) with high accuracy. This represents the "*current model state*". The points in the "*Unknown-knowns*" category are not correctly predicted by the ML model although the model is sure about it. This is bound to model performance and can be improved by "*transfer learning*" or a better model design. The third category is "*Known-Unknown*". The ML model knows that these data points are confusing. For instance, the model cannot confidently tell whether these news articles are relevant or irrelevant. These data points are close to the decision boundaries of the model. Such data points are queried by "*uncertainty sampling*". "*Unknown-unknown*" data is beyond the model's knowledge since they are not labeled yet and could change the perspective of the model when labeled. This is caused by lacking diversity of data and can be tackled by "*diversity sampling*". The model thinks it is sure about such news but has not got a labeled sample from this theme which could lead to false prediction.

3.1 Uncertainty Sampling

Classification models try to draw a decision boundary between labeled data points, which allows them to distinguish and assign these data points as shown in Figure 2.a. The data points are again projected to 2D via TSNE [18] just for better illustration. In the Figure, the model is trained based on labeled data (yes/no) and the decision boundary is illustrated by the hypothetical (manually drawn) line. In this example case, the model predicts points on the left and right of the boundary as relevant (yes) and irrelevant (no) respectively. The model tends to predict the points away from the decision boundary confidently. However, the points closer to the decision boundary are risky for the model to predict. Uncertainty sampling aims to query such points. In the below example, the points queried by the uncertainty sampling strategy are in the region of the boundary as expected. Uncertainty query strategies try to select data for labeling which could contribute to the model performance the most [16]. Thus, after labeling each query, the model can distinguish categories better which improves accuracy. For instance, the environment context model becomes better in distinguishing the relevancy of known themes after each uncertainty sampling query. However, what if the labeled data only consists of certain topics? Then the model assumes only these topics exist and tries to classify only them. Therefore, the variety of data is vital, which is achieved via diversity sampling.



(a) (b)

Figure 2: Queries taken via simulation based on true labels starting from the same state for ten iterations with different strategies (a) Uncertainty Sampling vs. (b) Diversity Sampling. A hypothetical decision boundary is manually drawn for uncertainty sampling to emphasize pattern. Data were taken from a BBC News dataset [17].

3.2 Diversity Sampling

Diversity sampling aims to query representative data points from the unlabeled data set [14] . The target is to solve the problem mentioned above. Uncertainty sampling without diversity sampling would focus more on the insights provided by the currently labeled data, and if it is not representative, the model will have gaps [15]. This phenomenon is essential for environment analysis since the expected result is to provide comprehensive information from the environment to the user as much as possible. If certain topics are overlooked and not labeled, the model would be trained without this knowledge and make inaccurate predictions. For example, if the labeled data contains only news about electronics, the uncertainty sampling tries to clarify the decision boundary for this topic (see Figure 2.a). In this case, topics such as Covid-19 or logistics which might be further away from electronics articles in vector space will be ignored. This will result in an incomplete recommendation from the context model. Diversity sampling helps to overcome this issue as shown in Figure 2.b. Nevertheless, diversity sampling will not focus on the data which might be predicted incorrectly by the model and this results in a less marginal positive impact on model performance compared to uncertainty sampling.

3.3 Hybrid and Advanced Strategies

The trade-off between diversity and uncertainty sampling strategies can be addressed by hybrid and combined models. Hybrid strategies merge both diversity and uncertainty into a single strategy [14]. Querying Informative and Representative Examples (QUIRE) [19], Learning Active Learning (LAL) [20], Self-Paced Active Learning (SPAL) [21] are some applications. However, these methods try to solve the problem end-to-end via complex metrics [14] which hinders interpretability for use cases and algorithms may become slower. This causes latency and more expert time for labeling. To address this issue, MUNRO [15] suggests advanced strategies which try to combine the advantages of two strategies more explicitly by applying them sequentially. For example, uncertainty sampling could be used to query some data points which could be further filtered by diversity sampling. However, this involves narrowing down samples via simple strategies one by one.

◔ Low
◑ Below Average
◕ Above Average
● High

		Query Strategies								
		Uncertainty Sampling		Diversity Sampling		Hybrid Strategies			Combined Strategies	
		Least Confidence / Margin of Confidence / Entropy [14–16]	Query-By-Committee [14,16]	Density-Based/ Representative Sampling [14,16]	Cluster-based Sampling [14,15]	QUIRE [19,22]	LAL [20,22]	SPAL [21,22]	Least Confidence with Cluster-based Sampling [15]	Representative Cluster-based Sampling [15]
Criteria	Variety of samples for complete topic coverage	◔	◔	◕	◕	◕	◑	◕	◑	●
	Contribution to model performance for better distinction of relevancy	●	●	◔	◔	◕	◕	◑	◕	◔
	Interpretability of metric used during query to check use case compatibility	●	◕	●	◕	◔	◔	◔	◕	◕
	Speed of queries to reduce latency and expert time spent	●	◕	●	◕	◔	◔	◑	◕	◕

Figure 3: Comparison of existing strategies based on designated criteria for environment analysis

Figure 3 provides the evaluation of strategies from each type based on the criteria essential for the environment analysis labeling procedure. It can be seen that combining strategies can supplement diversity and uncertainty simultaneously at the cost of speed and simplicity.

4. Application

AL query strategy categories for classification problems and the trade-off between them are explained in the previous section. The context model for environment analysis solves a text classification problem and AL is a suitable method because oracle (i.e. expert, user) can label the data intuitively. However, expert time for labeling should be minimized and the labeled dataset should be adequately diverse and informative for the model. Moreover, the labeling experience for the user should be smooth. Considering all these aspects, a web-based prototype for AL with a modular backend for different strategies is developed and actively used for labeling by the industry partners. Python programming language is used for development.

4.1 User Interface

AL queries data based on the current state of the initial (labeled) and pool (unlabeled) dataset and learning algorithm for data selection should be regularly updated after each iteration of labeling, especially for uncertainty sampling. Therefore, it is not possible to give a static datasheet to a user to label data. There are AL strategies that enable batch sampling. These are especially valuable if the query strategy is time-consuming and computationally heavy, but they are likely to suffer from adaptivity and redundancy [14]. Therefore, they are not preferred especially for uncertainty sampling.

Ever Given: Ship that blocked Suez Canal sets sail after deal signed

A huge container ship that blocked the Suez Canal in March - disrupting global trade - is finally leaving the waterway after Egypt signed a compensation deal with its owners and insurers.The Ever Given weighed anchor shortly after 11:30 local time (09:30 GMT) and headed north towards the Mediterranean escorted by tugs.

Source: https://www.bbc.com/news/world-middle-east-57746424

Total: 1490 | Labeled: 77 | Unlabeled: 1413 | Labeled Ratio: 5.2%

Labeled Data (Page Size = 50)

	Text	Label
	filter data...	
○	format wars could confuse users technology firms sony philips matsushita and samsung are developing a common way to stop people pirating digital	yes
○	smart search lets art fans browse if you don t know art but know what you like new search technology could prove a useful gateway to painting.	yes
○	uk gets official virus alert site a rapid alerting service that tells home computer users about serious internet security problems is being launched	yes
○	fockers fuel festive film chart comedy meet the fockers topped the festive box office in north america setting a new record for christmas day. the	no
○	housewives lift channel 4 ratings the debut of us television hit desperate housewives has helped lift channel 4 s january audience share by 12%	no
○	lewis-francis eyeing world gold mark lewis-francis says his olympic success has made him determined to bag world championship 100m gold in 2005. the	no
○	britain boosted by holmes double athletics fans endured a year of mixed emotions in 2004 as stunning victories went hand-in-hand with disappointing	no
○	worldcom ex-boss launches defence lawyers defending former worldcom chief bernie ebbers against a battery of fraud charges have called a company	no
○	greek pair set for hearing kostas kenteris and katerina thanou will fight the provisional two-year bans imposed on them by the iaaf at an independent	no
○	obituary: dame alicia markova dame alicia markova who has died in bath aged 94 was the uk s first prima ballerina of the modern age and in her heyday	ignore
○	wales silent on grand slam talk rhys williams says wales are still not thinking of winning the grand slam despite a third six nations win. that s	no
○	ultimate game award for doom 3 sci-fi shooter doom 3 has blasted away the	yes

Figure 4: Active Learning Interface

A web-based user interface is developed to provide a smooth and adaptive labeling experience (Figure 4). The learning algorithms and strategies are updated in the backend when necessary while displaying selected queries in the front end. The current news headline and its description are shown on the top left with three buttons for labeling: relevant (blue), not relevant (yellow), ignore/not sure (gray). Ignore option is included to reduce noisy labels from users when they are not confident and they are simply ignored. Such data points are no longer queried for labeling and are not used to train the model since the user is not sure about them. On the bottom left, a paginated list of labeled data (history) can be observed, filtered, selected and modified. The articles on the list are displayed by an accompanying scatter plot where they are projected on the 2D via dimensionality reduction (TSNE[18]) and colored in line with labeling buttons. The models are trained on high dimensional vectors, but reduced 2D coordinates are shown just for interactivity. Moreover, these compressed/reduced 2D vectors still contain patterns in original data, so users can see forming clusters. On the top right, the current status can be observed. This simple interface allows users to label the data quickly while getting adaptive queries from the backend as described in Figure 1.

4.2 Modular Backend: Chaining Strategies

The aim of the modular backend is plain. The query strategies are chained one by one in order. Active strategy selects instances and the user labels them. After the designated number of queries are made with the strategy, the next strategy is initiated. The idea can be described as follows. The chain comprises diversity and uncertainty sampling methods interchangeably. Diversity sampling tries to fill the gaps in the information available to the model by selecting representative instances based on the current state. This is expected to confuse the model because new topics or themes will be introduced. In other words, the decision boundaries in some areas of feature space become vague. In the next step, uncertainty sampling strategies try to reduce disruption caused by increased diversity. This is done by selecting points to improve the model's performance. This exchange between strategies is repeated as much as necessary. Active Learning strategies try to maximize marginal contribution on accuracy as covered in the literature on individual strategies [23,19,24–26]. The exchange between diversity and uncertainty sampling improves coverage on topics in the dataset and distinction capability on these topics respectively.

The modular structure is developed based on the modAL (Modular Active Learning) framework [27]. It contains some strategies by default, but it also provides plug-and-play option for custom-made query strategies. Strategies such as basic uncertainty sampling, margin sampling, entropy sampling and query by committee are inherited from default functions [27]. Nevertheless, strategies such as cluster-based AL [15], density-based sampling [14] or advanced strategies like uncertainty-representative sampling [15,16] are developed and implemented. Moreover, these strategies are optimized for low latency sampling. Custom and tailored methods are also developed. One strategy tries to boost the number of positively labeled data (relevant) to reduce unbalance. Another strategy aims to increase diversity based on Radial Basis Function (RBF) Kernel [28] based similarity. They are also implemented as a strategy in the modular chain.

Order, repetition and parameters for strategies can be set by a config file. If there is no initial data, the chain starts with a cluster-based or kernel-based strategy since they do not require a pre-trained model on labeled data to query. Moreover, the query with this initial strategy continues until the minimum required amount of data from each category (relevant-not relevant) is obtained. Then, the following strategies in the chain are activated in order. The latency is reduced since algorithms are optimized for speed or selected accordingly. Preprocessing of pool (unlabeled) data including vectorization, clustering and dimensionality reduction is also done beforehand to prevent unnecessary computations during deployment. The combination of these properties, chained strategies and interactive interface provides a seamless experience for the user to label the informed selection of samples.

5. Discussion

The modular structure uses individual strategies, which are not novel by themselves and most of them are implemented based on existing methods. The main objective or contribution of this presented solution is to utilize the labeling effort of the user most efficiently. A context model will be the output of labeled data. This context model should mimic the perspective and mindset of the users as much as possible, to filter new information based on their preferences.

The current AL solution by nature provides intelligence to the labeling process. However, manual chaining of strategies via a config file can be improved. The goal sought by each strategy can be defined. Then, the current state during the labeling can be monitored via metrics or algorithms. These metrics could be used to trigger the transition between strategies. For example, if model performance is no longer improved by the current iteration of uncertainty sampling, diversity sampling can be initiated. The other aspect to consider is the static pool of data used during the labeling process, even if the data is up-to-date at that time. The context model will be trained based on this data. Therefore, methods should be implemented to recommend new topics to the user and learn their preferences toward them. This can be achieved by integrating trend detection methods [29] or tools [30] in the pipeline.

6. Summary and outlook

Awareness and vigilance are valuable in today's turbulent business environment. The utilization of big data could help to reap this benefit. Environment analysis tries to find the relevant information for an organization from the outside world and present it in a structured manner. This paper presents an approach to develop the first step of the pipeline for this analysis. The necessity of labeled data is explained, and the application of Active Learning is shown to satisfy this prerequisite efficiently and seamlessly. The goal is to minimize the expert time spent and amount of data labeled as much as possible to have a comprehensive understanding of which information the corresponding organization deems worthy. The user interface and modular backend consisting of chained query strategies, some of which are optimized and tailored for the problem, enable the goal. The overcoming of limitations originated from lack of intelligent transition between query strategies and integration of new trends and topics into the finalized system should be considered. The next step is to filter and analyze the data further based on the integrated enterprise model (IEM) [31] and the data of enterprise IT systems from an industry partner. Finally, relevant and tailored information for the organization will be displayed on the situation awareness monitor [5].

Acknowledgements

This research and development project is funded by the German Federal Ministry of Education and Research (BMBF) within the "Innovations for Tomorrow's Production, Services, and Work" Program (02P18X000) and implemented by the Project Management Agency Karlsruhe (PTKA). The author is responsible for the contents of this publication.

References

[1] Vecchiato, R., 2015. Creating value through foresight: First mover advantages and strategic agility. Technological Forecasting and Social Change 101, 25–36.

[2] Rust, H., 2021. Weise Voraussicht und Erfolgsplanung. Springer Fachmedien Wiesbaden, Wiesbaden, 210 pp.

[3] Dadkhah, S., Bayat, R., Fazli, S., Tork, E.K., Ebrahimi, A., 2018. Corporate foresight: developing a process model. Eur J Futures Res 6 (1), 1–10.

[4] Jain, S., 1984. Environmental Scanning in U.S. Corporations. Long Range Planning (Vol. 17, No. 2), 117–128.

[5] Kohl, H., Knothe, T., Oertwig, N., Gering, P., Scholz, J.-A., 2021. Interaktives Lagebild: Ein Werkzeug für das Krisenmanagement in prozessorientieren Unternehmen. Industrie 4.0 Management 37 (1), 37–40.

[6] Will, M., 2008. Talking about the future within an SME?: Corporate foresight and the potential contributions to sustainable development. Management of Env Quality 19 (2), 234–242.

[7] IBM Cloud Education, 2020. Natural Language Processing (NLP). https://www.ibm.com/cloud/learn/natural-language-processing. Accessed 15 March 2022.

[8] IBM Cloud Education, 2021. Structured vs. Unstructured Data: What's the Difference? https://www.ibm.com/cloud/blog/structured-vs-unstructured-data. Accessed 5 January 2022.

[9] BBC, 2021. Egypt's Suez Canal blocked by huge container ship. https://www.bbc.com/news/world-middle-east-56505413. Accessed 5 January 2022.

[10]Elasticsearch B.V., 2021. What is Elasticsearch? | Elasticsearch Guide [7.16] | Elastic. https://www.elastic.co/guide/en/elasticsearch/reference/current/elasticsearch-intro.html. Accessed 5 January 2022.

[11]Delua, J., 2021. Supervised vs. Unsupervised Learning: What's the Difference?, March 12.

[12]scikit-learn developers, 2022. 2.3. Clustering. https://scikit-learn.org/stable/modules/clustering.html. Accessed 5 January 2022.

[13]Salas-Eljatib, C., Fuentes-Ramirez, A., Gregoire, T.G., Altamirano, A., Yaitul, V., 2018. A study on the effects of unbalanced data when fitting logistic regression models in ecology. Ecological Indicators 85, 502–508.

[14]Kumar, P., Gupta, A., 2020. Active Learning Query Strategies for Classification, Regression, and Clustering: A Survey. Journal of Computer Science and Technology 35 (4), 913–945.

[15]Munro, R., 2021. Human-in-the-Loop Machine Learning, 1st edition ed. Manning Publications; Safari, Erscheinungsort nicht ermittelbar, Boston, MA, 424 pp.

[16]Settles, B. Active Learning Literature Survey. University of Wisconsin-Madison Department of Computer Sciences. https://minds.wisconsin.edu/handle/1793/60660.

[17]Greene, D., Cunningham, P., 2006. Practical Solutions to the Problem of Diagonal Dominance in Kernel Document Clustering, in: Proc. 23rd International Conference on Machine learning (ICML'06). ACM Press, pp. 377–384.

[18]scikit-learn developers, 2022. TSNE. https://scikit-learn.org/stable/modules/generated/sklearn.manifold.TSNE.html. Accessed 14 March 2022.

[19]Huang, S.-J., Jin, R., Zhou, Z.-H., 2014. Active learning by querying informative and representative examples. IEEE Transactions on Pattern Analysis and Machine Intelligence 36 (10), 1936–1949.

[20]Konyushkova, K., Sznitman, R., Fua, P., 2017. Learning Active Learning from Data. Advances in Neural Information Processing Systems 30.

[21]Tang, Y.-P., Huang, S.-J., 2019. Self-Paced Active Learning: Query the Right Thing at the Right Time. AAAI 33, 5117–5124.

[22]Tang, Y.-P., Li, G.-X., Huang, S.-J., 2019. ALiPy: Active Learning in Python. https://arxiv.org/pdf/1901.03802.

[23]Du, B., Wang, Z., Zhang, L., Zhang, L., Liu, W., Shen, J., Tao, D., 2015. Exploring representativeness and informativeness for active learning. IEEE transactions on cybernetics 47 (1), 14–26.

[24]Krempl, G., Kottke, D., Spiliopoulou, M. Probabilistic active learning: Towards combining versatility, optimality and efficiency, in: International Conference on Discovery Science. Springer, pp. 168–179.

[25]Lewis, D.D., Gale, W.A., 1994. A Sequential Algorithm for Training Text Classifiers. https://arxiv.org/pdf/cmp-lg/9407020.

[26]Nguyen, V.-L., Shaker, M.H., Hüllermeier, E., 2022. How to measure uncertainty in uncertainty sampling for active learning. Mach Learn 111 (1), 89–122.

[27]Danka, T., Horvath, P., 2018. modAL: A modular active learning framework for Python, 5 pp. http://arxiv.org/pdf/1805.00979v2.

[28]scikit-learn developers, 2022. RBF. https://scikit-learn.org/stable/modules/generated/sklearn.gaussian_process.kernels.RBF.html. Accessed 15 March 2022.

[29]Kontostathis, A., Galitsky, L.M., Pottenger, W.M., Roy, S., Phelps, D.J., 2004. A Survey of Emerging Trend Detection in Textual Data Mining, in: Berry, M.W. (Ed.), Survey of Text Mining. Clustering, Classification, and Retrieval. Springer, New York, NY, pp. 185–224.

[30]Synthesio, 2013. Trend Analysis Tools | Synthesio. https://www.synthesio.com/glossary/trend-analysis-tools/. Accessed 5 January 2022.

[31]Spur, G., Mertins, K., Jochem, R., 1993. Integrierte Unternehmensmodellierung, 1. Aufl. ed. Beuth, Berlin.

Biography

Furkan Agacayaklar (*1993) is a research associate in the business process and factory management department at the Fraunhofer Institute for Production Systems and Design Technology (IPK) Berlin since 2021 with focus on data applications. He has previously studied Mechanical Engineering (B.Sc.) at Bogazici University and Management and Engineering in Production Systems (M.Sc.) at RWTH Aachen.

Annika Lange (*1996) is a research associate in the business process and factory management department at the Fraunhofer Institute for Production Systems and Design Technology (IPK) Berlin since 2021. She has previously studied Sports Engineering at Otto-von-Guericke University Magdeburg (B.Sc.) with an exchange at the University college of southeast Norway and Mechanical Engineering at Leibniz University Hannover (M.Sc.).

Julia-Anne Scholz (*1992) is a research associate at the Fraunhofer Institute for Production Systems and Design Technology (IPK) and has been developing solutions for agile process management and resilient business alignment. She has previously studied Industrial Engineering at the Technical University of Berlin with an exchange at the Beijing Tsinghua University.

Prof. Dr.-Ing. Thomas Knothe (*1971) has been head of the Business Process and Factory Management Department at the Fraunhofer Institute for Production Systems and Design Technology (IPK) Berlin since 2010. He has previously studied Information Technology in Mechanical Engineering at the Technical University of Berlin and completed his doctorate in engineering.

Dirk Busse (*1969) is founder, shareholder and managing director of budatec GmbH since 2009. He previously made an apprenticeship as Precision Mechanic and studied Mechanical Engineering at the Technical University Berlin (Dipl-Ing.).

3rd Conference on Production Systems and Logistics

Review Of Blockchain-based Tokenization Solutions For Assets In Supply Chains

Fabian Dietrich[1,2], Nils Kuenster[1], Louis Louw[2], Daniel Palm[1,3]
[1] *ESB Business School, Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany*
[2] *Department of Industrial Engineering, Stellenbosch University, 145 Banghoek Rd., 7600, Stellenbosch, South Africa*
[3] *Fraunhofer Institute for Manufacturing Engineering and Automation, Alteburgstr. 150, 72762 Reutlingen, Germany*

Abstract

Recently, blockchain-based tokens have earned an important role in fields such as the art market or online gaming. First approaches exist, which adopt the potentials of blockchain tokens in supply chain management to increase transparency, visibility, automation, and disintermediation of supply chains. In context, the tokenization of assets in supply chains refers to the practice of creating virtual representations of physical assets on the blockchain. Solutions in supply chain management based on the tokenization of assets vary in terms of application objectives, token types, asset characteristics, as well as the complexities of supply chain events to be mapped on the blockchain. Currently, however, no review exists that summarizes the characteristics of blockchain-based tokens and their scope of applications. This paper provides a clear terminological distinction of existing blockchain token types and therefore distinguishes between fungible tokens, non-fungible tokens, smart non-fungible tokens, and dynamic smart non-fungible tokens. Subsequently, the token types are classified regarding their traceability, modifiability, and authorization to evaluate suitability for mapping assets in supply chains. Given the potential of blockchain in supply chain management, the results of the review serve as a foundation for a practical guide supporting the selection process of suitable token types for industrial applications.

Keywords

Supply Chain Management; Blockchain; Tokenization; Traceability; Review

1. Introduction

Blockchain-based tokens have significantly grown in popularity and public attention. Especially so-called Non-Fungible Tokens (NFTs) have created a growing digital market for artworks, collectibles, and other digital assets. In recent years the development shows a shift from digital markets almost exclusively trading digital art to increasingly different categories such as items in games or metaverses [1]. Furthermore, there are emerging applications aiming to tackle real-world problems using NFTs. Creating and selling NFTs could post a viable option to raise financing and awareness for art museums or wildlife conservation efforts [2,3]. Other applications include the management of privacy when sharing genetic data with health care providers to mitigate the risk of sensitive information abuse [4]. The concept of NFTs is also progressively being adapted in an industrial context. In particular, blockchain with its NFTs bears the potential to increase transparency, automation, visibility, and disintermediation in Industry 4.0 driven supply chains [5].

The idiosyncrasy of blockchain tokens results from the technical and organizational properties of blockchain technology. In 2008, the pseudonym Satoshi Nakamoto published the Bitcoin white paper and thus

DOI: https://doi.org/10.15488/12147

introduced blockchain technology with the aim of creating a secure electronic currency independent of trusted third parties [6]. The US National Institute of Standards and Technology defines blockchain technology as “distributed digital ledgers of cryptographically signed transactions that are grouped into blocks. Each block is cryptographically linked to the previous one (making it tamper evident) after validation and undergoing a consensus decision” [7]. The emergence of Ethereum, introduced in Vitalik Buterin’s white paper in 2013 [8], extended the scope of blockchain from financial applications to other fields. The Ethereum white paper marks the start of blockchain-based ‘smart contracts’ as an enabler for further decentralized applications. Ethereum integrates the characteristics of blockchain with a fully-fledged Turing-complete programming language. Following the mathematical concept of Turing-completeness, which inter alia provides means to compute loops and/or complex recursions [9], Ethereum represents the first general-purpose blockchain platform that is able to compute and run all kinds of complex logical constructs, which serves as the technological foundation for the so-called ‘tokenization of assets’ [10].

2. Related works

The process of creating virtual representations of physical assets on the blockchain is usually referred to as the ‘tokenization of assets’ [11]. Initially, supply chain assets were tokenized to prove the ownership and provenance of the underlying physical asset [12]. Evolving blockchain technology and platforms allowed the extensions of blockchain-based tokenization to supply chain transparency applications [13]. Here, the definition and adaption of token standards played an important role in facilitating the integration of tokens for such applications. Currently, the Ethereum blockchain still implements the most pertinent token standards [14]. The establishment of the Ethereum standards ERC-20 and ERC-721 led to the common distinction between different types of blockchain tokens, namely Fungible Tokens (FTs) and the already mentioned NFTs. Different units of token that conform to ERC-20 can be interchanged with each other and therefore are, as the term suggests, fungible. The ERC-721 standard on the other hand implements tokens with unique identifications making tokens non-fungible [14]. On the Ethereum blockchain, the multi-token standard ERC-1155 extends the functionalities of the ERC-721 standard with the possibility of deploying several tokens using one smart contract [15].

These token standards, however, only represent minimum specifications of required functions to implement certain applications. The functionalities of a token itself are thereby determined by the underlying smart contract. Extensions of these standards are necessary to address the challenges of more complex structures such as supply chain applications. Arcenegui et al. [16] describe tokens that require an extension regarding their core functionalities as ‘Smart NFTs’.

Westerkamp et al. [17] present an extension to the ERC-721 token standard specifically aiming at solving the challenges of manufacturing supply chains. To map complex manufacturing processes that include the assembling of parts the authors introduce so-called ‘Token Recipes’. This enables the tracing of assembled goods and their components throughout supply chains [17]. Elaborating on this idea Watanabe et al. [18] introduce a token structure that embeds a link to previous states in the token transactions [18]. This enables more efficient retrieval of information in blockchain networks, which facilitates the tracing capability of tokens and therefore preferably suits applications aiming at increasing supply chain transparency. Kuhn et al. [19] present an approach utilizing the Ethereum multi-token standard ERC-1155 to build tokens mapping complex assembly structures. The ERC-1155 standard includes the possibility to deploy multiple fungible and non-fungible tokens using a single smart contract [15]. The approach of Kuhn et al. extends this smart contract to a so-called ‘Assembly Token Manager’ [19]. This manager governs the token balances and transformation events while also providing traceability information through an event log. To increase the viability of token-based applications in dynamic environments, Dietrich et al. [20] present a token concept allowing the definition and assignment of clear authorities when deploying the smart contract. These authorities can consent to dynamic changes on a deployed token smart contract such as adding new parts or partners to the supply chain. Additionally, this approach embeds the token history and composition into the token structure making external event logs superfluous. This streamlines the ability to holistically map complex and dynamic supply chains.

As the breakdown of the different approaches exemplifies, literature shows a range of different token solutions tackling the adaptation of blockchain tokens for supply chain applications. With the adaptation of blockchain tokens for increasingly complex application scenarios, an incrementally more complex token landscape has emerged that extends far beyond the originally defined token standards. There is, however, a degree of overlap between the different token solutions as well as a lack of uniform terminology. Therefore, this review of tokenization solutions aims on the one hand to provide a clear terminological distinction of existing blockchain token types and on the other hand to serve as a foundation for a practical guide supporting the selection process of suitable token types for industrial applications.

3. Review of tokenization solutions

The following chapter comprises a review of blockchain-based tokenization solutions for assets in supply chains. The first section defines the dimensions of classifiable differences by investigating the process of tokenization in comparison with already established identification systems. The second section examines the extent to which the complexity of the underlying asset affects the mapping by means of tokenization and derives important token characteristics. Lastly, a token classification is presented considering different token designs of existing approaches.

3.1 Tokenizing of assets

For the tokenization of supply chain assets, it is necessary to create a virtual reflection of assets on the blockchain [10]. This requires the introduction of an asset-backed token on the blockchain as well as a clear linkage of the virtual token to the physical or abstract real-world asset. In the case of physical assets, the linkage can be realized by using identification technologies such as RFID or QR Codes [21,22]. Such identification systems are subject to extensive legal and technical standardization. The IEC 62507-1:2010 standard provides a reference model for the reflection of assets from the physical world into the virtual world, which includes the combination of metadata, a unique identifier, and a physical or abstract asset [23].

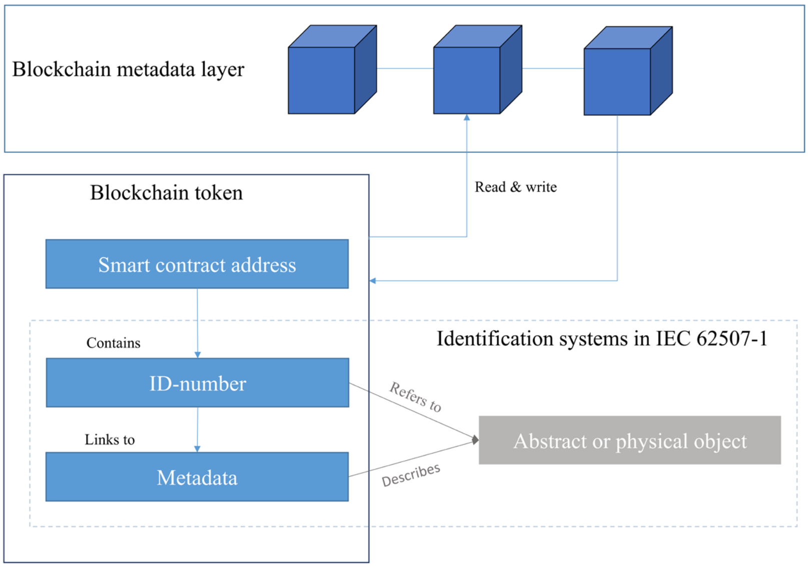


Figure 1: Referencing model for tokenization of assets

Here, the unique identifier refers to the physical or abstract object and links to its metadata. Accordingly, an object's metadata describes the physical or abstract object. The complexity of tokenization, however, goes beyond the IEC 62507-1:2010 reference model, in particular since NFTs eventually represent unique assets in publicly available networks. Figure 1 shows an extended reference model specifically designed according to the properties of NFTs. The composition of an NFT is determined by a token smart contract, which can be identified with its unique address on the blockchain. The smart contract contains the unique token identifier, typically represented by hexadecimal numbers as a unique result of underlying hashing functions [17,20]. This identifier links to the token's metadata. As Figure 1 indicates, the connection between the unique identifier and the metadata overlaps with the reference model of the IEC 62507-1:2010 standard. Furthermore, the extended model takes the relationship between the blockchain metadata layer into account. Typically, smart contracts embedded into a blockchains block structure own the ability to exchange information with the metadata layer itself. The extended reference model only describes the relationships between the aspects of tokenization. However, it does not specify to what extent data and functions must be included in the token architecture in order to meet the requirements of the respective application scenario to be mapped on the blockchain.

3.2 Adapting tokens to asset complexity

To enable the mapping of supply chain assets by means of tokenization, blockchain tokens must meet the requirements of the assets as well as the supply chain network. This involves the mapping of information and material flows as well as relationships within the supply chain network [24]. To facilitate such mapping, information systems rely on a standardized division into so-called supply chain events, which determine "what, when, where and why" something happens [25]. Previous research has already shown that the underlying supply chain complexity in terms of different supply chain events to be mapped impacts the requirements and selection of different token types [5]. In this context, Bozarth et al. [26] distinguish between the detail complexity, which describes the "distinct number of components or parts that make up a system", and the dynamic complexity, which deals with "the unpredictability of a system's response to a given set of inputs, driven in part by the interconnectedness of the many parts that make up the system". Derived from the intended application of blockchain tokens and the requirements arising from an increased supply chain complexity, this paper identifies three characteristics impacting the complexity of assets to be mapped on the blockchain.

Traceability: Traceability can be described as the ability to gather information on the whole downstream path of a product [27]. Companies have to deal with the growing interests of customers, governments, and non-governmental organizations in having greater transparency of brands, manufacturers, and producers throughout the supply chain [28,29]. As a result, social and environmental sustainability issues have become increasingly important for manufacturers in order to maintain the flawless reputation of their brands [30]. The ability to mitigate the risks of counterfeit products and the enforcement of sustainability standards in multi-tier supply chains rely on the effective traceability of assets [31]. Furthermore, emerging legal requirements increase the pressure on companies to ensure traceability throughout their supply chains. For example, initiatives such as the "EU rules on due diligence in supply chains" aim at making companies accountable for their entire supply chain [32].

Modifiability: Modifiability refers to the ability of an asset to experience transformation and aggregation events throughout its lifecycle according to the EPICS standard by GS1 [25]. This includes assets to eventually experience changes in terms of their modular composition. The modularity of assets describes the division of a structure into different subsections that can be combined using interfaces and may be replaced or reused in other products [33]. While this practice can reduce the overall complexity of an asset's structure it increases the number of processing events that can occur in a given supply chain. The modularization of products represents an important tool in achieving mass customization capabilities [34,35]. The aim to deliver customized products at the cost of mass production is closely tied to the emergence of the ideas of Industry 4.0 [36]. Furthermore, modular product architectures along with the ability to disassemble and reuse the modules are central characteristics of products in a circular economy [37].

Authorization: The collaboration in supply chain networks requires the establishment of authorization mechanisms governing the role and activities of participating supply chain members [38]. The sweeping collaboration between partners in supply chains has the potential to create benefits for every partner and ultimately a competitive advantage for the whole supply chain [39]. The need for collaboration is driven by the occurrence of uncertainties, disruptions, and dynamic changes in modern supply chains [40,41].

3.3 Token type classification

When summarizing the existing approaches adopting supply chain tokens and investigating their properties in terms of traceability, modularity, and authorization, four main categories of token types emerge.

These token types, Fungible Token, Non-Fungible Token, Smart Non-Fungible Token, and Dynamic Smart Token are classified in Table 1 according to their functional and data properties in terms of traceability, modularity, and authorization.

Table 1: Token type classification

			Token types			
			Fungible Token	**Non-Fungible Token**	**Smart Non-Fungible Token**	**Dynamic Smart Non-Fungible Token**
Token characteristics	**Token traceability**	**Functionality**	Not possible	Possible	Possible	Possible
		Data	Not available	Retrievable from blockchain metadata layer: • Transaction history	Retrievable from blockchain metadata layer: • Transaction history • State history	Retrievable from token: • Transaction history • State history
	Token modifiability	**Functionality**	Not possible	Not possible	Possible events: • Transformation • Aggregation	Possible events: • Transformation • Aggregation • Disaggregation
		Data	Not available	Not available	Event history retrievable from blockchain metadata layer	Event history logically coupled to the token
	Token authorization	**Functionality**	Not possible	Possible verification: • Token authenticity	Possible verification: • Token authenticity • Account and role authenticity	Possible verification: • Token authenticity • Account and role authenticity • Administrative permissions
		Data	Not available	Accessible via blockchain metadata layer: • Smart contract address	Accessible via blockchain metadata layer and embedded token functions: • Smart contract address • Account and role data	Accessible via blockchain metadata layer and embedded token functions: • Smart contract address • Account and role data • Permissions in authorization concept

Fungible Tokens: Fungible Tokens are closely resembled by the specifications of the ERC-20 token standard [42]. They facilitate in particular the mapping of volume exchanges of a given good. However, they do not offer technical properties to enable the tracing of individual tokens as well as the implementation of a smart contract logic to govern token modifications and interactions. One conceivable scenario for the use of fungible tokens in a supply chain context is the token-based exchange of freight pallets, where only the total volume and available quantity are relevant for traceability purposes.

Non-Fungible Tokens: NFTs are closely resembled by the specifications of the ERC-721 token standard [43]. NFTs possess unique identifiers and therefore allow linking tokens to a physical or abstract asset according to the extended reference model shown in Figure 1. This enables token traceability by extracting the transaction history from the blockchain metadata layer. Pure NFTs do not offer the possibility to add extensive creation requirements to the token smart contract allowing mapping of the modularity of assets. Furthermore, it is only possible to verify a token's authenticity without connecting it to governance models. In a supply chain context, a conceivable scenario for applying NFTs is the management of tools that allows a flexible distribution and tracking of tools across a supply chain network.

Smart Non-Fungible Tokens: As Arcenegui et al. [16] describe, Smart NFTs extend the core functionalities of the NFT standards with supply chain specific functions and creation requirements. While the core properties regarding the traceability remain the same as with NFTs, Smart NFTs allow users to include token transformation and aggregation functions. These functions can additionally be assigned to first authoritative permissions. Thus, Smart NFTs allow the mapping of assets that experience predictable changes regarding their modular composition throughout the supply chain, such as comprehensively certified medical devices.

Dynamic Smart Non-Fungible Tokens: Dynamic Smart NFTs extend the idea of Smart NFTs by adding dynamic elements to the functions and creation requirements as well as embedding an authority concept into the token smart contract. Furthermore, Dynamic Smart NFTs enable the inclusion of a token's history and composition inside their token structure [20]. This functionality forms the basis for enabling not only an aggregation of tokens but also a subsequent disaggregation. A conceivable scenario for Dynamic Smart NFTs is the mapping of assets that can experience dynamic changes regarding their modular composition as well as underlying supply chain authority structure throughout the entire lifecycle of an asset, such as in the automotive industry.

4. Result

The review of different tokenization solutions of assets in supply chains results in a procedure for adopting blockchain tokens, which incorporates the modeling of tokens as well the classification of token types. Figure 2 shows the structured flow scheme of the procedure.

The first layer describes the necessity to initially define a clear scope of the supply chain and the respective assets to be mapped. This step must always be oriented towards the objective of the corresponding use case. Based on this, the second layer describes the adaption of token requirements according to the asset complexity. This includes an asset's requirements in terms of traceability, modifiability, and authorization as well as the respective functionalities and data. According to the derived requirements, a suitable token type can be selected in the last layer. Here, the flow scheme distinguishes between FTs, NFTs, Smart NFTs, and Dynamic Smart NFTs.

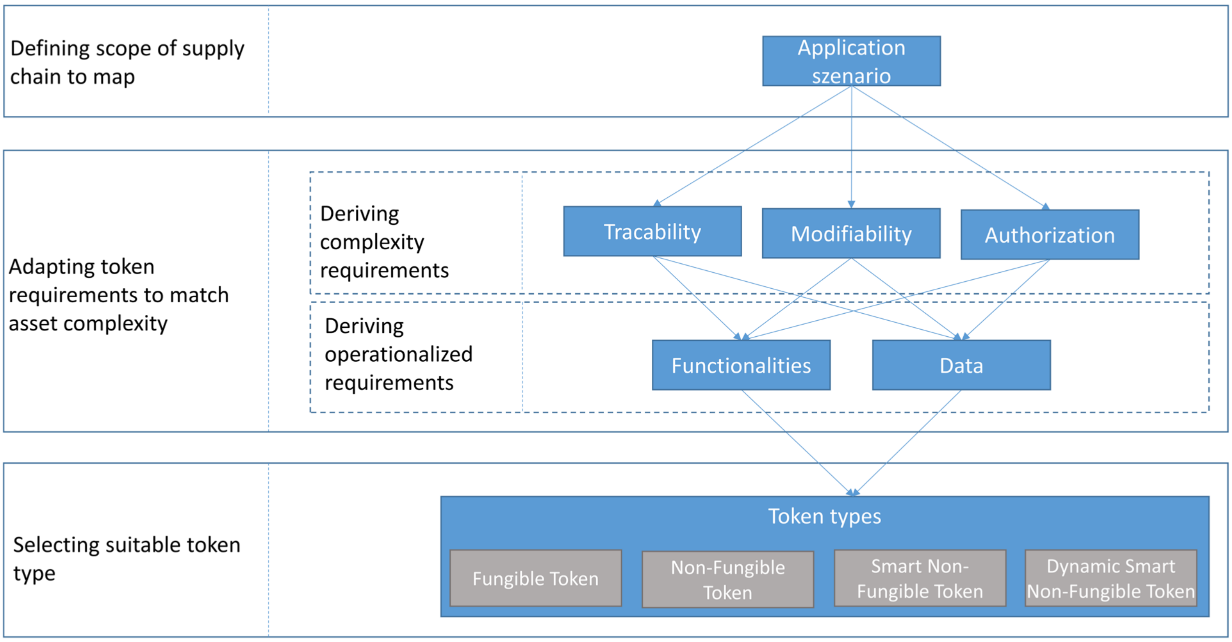


Figure 2: Procedure for adopting blockchain-based tokenization for assets in supply chains

5. Discussion and conclusion

The tokenization of assets is an innovative solution for mapping physical or abstract objects on the blockchain. This trend is also increasingly becoming important for industrial use cases bearing the potential to increase transparency, automation, visibility, and disintermediation in supply chains. The transfer of the tokenization of assets to increasingly complex use cases has led to the continuous development of different token solutions, in order to meet the increasing requirements regarding traceability, modifiability, and authorization. This paper provides a clear terminological distinction of existing blockchain token types and therefore distinguishes between FTs, NFTs, Smart NFTs, and Dynamic Smart NFTs. FTs and NFTs are very much based on the characteristics of the well-established token standards ERC-20 and ERC-721. Smart NFTs extend the token functions with supply chain specific requirements, which enables a static mapping of modifiable assets. Dynamic smart NFTs, as the name indicates, embed the token functions into dynamic authority and token concepts allowing the mapping of flexible supply chains with changeable assets. The findings are summarized in a procedure, which supports the selection process of suitable token types for industrial applications. So far, the token types classified in this paper have only been evaluated theoretically or based on a very limited amount of industrial case studies. This bears the potential that new token solutions with more extensive functionalities are necessary in order to meet industrial requirements holistically, which would require an extension of the available tokenization types. Therefore, further research is necessary to identify further requirements of assets to be mapped on the blockchain in a wide range of industrial domains. Currently, further research is being conducted in developing a framework serving as a practical implementation guide for industrial token-based applications.

References

[1] Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L.M., Baronchelli, A., 2021. Mapping the NFT revolution: market trends, trade networks, and visual features. Scientific Reports 11 (1), 20902.
[2] Valeonti, F., Bikakis, A., Terras, M., Speed, C., Hudson-Smith, A., Chalkias, K., 2021. Crypto Collectibles, Museum Funding and OpenGLAM: Challenges, Opportunities and the Potential of Non-Fungible Tokens (NFTs).
[3] Mofokeng, M., Matima, T., 2018. Future tourism trends: Utilizing non-fungible tokens to aid wildlife conservation. African Journal of Hospitality, Tourism and Leisure 7.

[4] Alghazwi, M., Turkmen, F., van der Velde, J., Karastoyanova, D., 2021. Blockchain for Genomics: A Systematic Literature Review. https://arxiv.org/pdf/2111.10153.

[5] Dietrich, F., Ge, Y., Turgut, A., Louw, L., Palm, D., 2021. Review and analysis of blockchain projects in supply chain management. Procedia Computer Science 180, 724–733.

[6] Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. www.bitcoin.org. https://bitcoin.org/bitcoin.pdf. Accessed 20 October 2019.

[7] Yaga, D., Mell, P., Roby, N., Scarfone, K., 2019. Blockchain technology overview. National Institute of Standards and Technology, Gaithersburg, MD.

[8] Buterin, V., 2013. Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform. https://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf. Accessed 20.10.19.

[9] Jansen, M., Hdhili, F., Gouiaa, R., Qasem, Z., 2020. Do Smart Contract Languages Need to Be Turing Complete?, in: Prieto, J., Das, A.K., Ferretti, S., Pinto, A., Corchado, J.M. (Eds.), Blockchain and Applications, vol. 1010. Springer International Publishing, Cham, pp. 19–26.

[10] Blossey, G., Eisenhardt, J., Hahn, G., 2019. Blockchain Technology in Supply Chain Management: An Application Perspective, in: Proceedings of the Annual Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences.

[11] de Filippi, P., Schuppli, B., Choi, C., Reyes, C., Divissenko, N., Lavayssière, X., Dagnino, F., Funke, M., Bierwirth,, F., Müller, T., et al., 2019. Regulatory Framework for Token Sales: An Overview of Relevant Laws and Regulations in Different Jurisdictions Coalition of Automated Legal Applications. Blockchain Research Institute.

[12] Van de Velde, J., Scott, A., Sartorius, K., Dalton, I., Shepherd, B., Allchin, C., Dougherty, M., Ryan, P., Rennick, E., 2016. Blockchain in Capital Markets - The Prize and the Journey. Oliver Wyman & Euroclear.

[13] Chang, S.E., Chen, Y.-C., Lu, M.-F., 2019. Supply chain re-engineering using blockchain technology: A case of smart contract based tracking process. Technological Forecasting and Social Change 144, 1–11.

[14] Antonopoulos, A.M., Wood, G.A., 2019. Mastering Ethereum - Building smart contracts and DApps. O'Reilly Media, Tokyo, 384 pp.

[15] Radomski, W., Cooke, A., Castonguay, P., Therien, J., Binet, E., Sandford, R., 2018. EIP-1155: Multi Token Standard. Ethereum Improvement Proposals.

[16] Arcenegui, J., Arjona, R., Román, R., Baturone, I., 2021. Secure Combination of IoT and Blockchain by Physically Binding IoT Devices to Smart Non-Fungible Tokens Using PUFs.

[17] Westerkamp, M., Victor, F., Küpper, A., 2020. Tracing manufacturing processes using blockchain-based token compositions. Digital Communications and Networks 6 (2), 167–176.

[18] Watanabe, H., Ishida, T., Ohashi, S., Fujimura, S., Nakadaira, A., Hidaka, K., Kishigami, J., 2019. Enhancing Blockchain Traceability with DAG-Based Tokens, in: 2019 IEEE International Conference on Blockchain (Blockchain). 2019 IEEE International Conference on Blockchain (Blockchain), pp. 220–227.

[19] Kuhn, M., Funk, F., Zhang, G., Franke, J., 2021. Blockchain-based application for the traceability of complex assembly structures. Journal of Manufacturing Systems 59, 617–630.

[20] Dietrich, F., Louw, L., Palm, D., 2022. Concept for a Token-Based Blockchain Architecture for Mapping Manufacturing Processes of Products with Changeable Configurations, in: Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems, Cham. 2022. Springer International Publishing, Cham, pp. 508–515.

[21] Tian, F., 2016. An agri-food supply chain traceability system for China based on RFID & blockchain technology, in: 2016 13th International Conference on Service Systems and Service Management (ICSSSM). 2016 13th International Conference on Service Systems and Service Management (ICSSSM), pp. 1–6.

[22] Toyoda, K., Mathiopoulos, P.T., Sasase, I., Ohtsuki, T., 2017. A Novel Blockchain-Based Product Ownership Management System (POMS) for Anti-Counterfeits in the Post Supply Chain. IEEE Access 5, 17465–17477.

[23] Identification systems enabling unambiguous information interchange - Requirements: Part 1: Principles and methods, 2010 IEC 62507-1:2010.

[24] Carter, C.R., Rogers, D.S., Choi, T.Y., 2015. Toward the Theory of the Supply Chain. J Supply Chain Manag 51 (2), 89–97.

[25] Kennedy, A., Troeger, R., Morgan, G., Traub, K., 2016. EPC Information Services (EPCIS) Standard. GS1. https://www.gs1.org/sites/default/files/docs/epc/EPCIS-Standard-1.2-r-2016-09-29.pdf.

[26] Bozarth, C.C., Warsing, D.P., Flynn, B.B., Flynn, E.J., 2009. The impact of supply chain complexity on manufacturing plant performance. Journal of Operations Management 27 (1), 78–93.

[27] Bosona, T., Gebresenbet, G., 2013. Food traceability as an integral part of logistics management in food and agricultural supply chain. Food Control 33 (1), 32–48.

[28] Chen, S., Zhang, Q., Zhou, Y.-P., 2019. Impact of Supply Chain Transparency on Sustainability under NGO Scrutiny. Prod Oper Manag 28 (12), 3002–3022.

[29] New, S., 2010. The Transparent Supply Chain. https://hbr.org/2010/10/the-transparent-supply-chain. Accessed 3 February 2022.

[30] Lemke, F., Petersen, H.L., 2018. Managing Reputational Risks in Supply Chains, in: Khojasteh, Y. (Ed.), Supply Chain Risk Management: Advanced Tools, Models, and Developments. Springer Singapore, Singapore, pp. 65–84.

[31] Hastig, G.M., Sodhi, M.S., 2020. Blockchain for Supply Chain Traceability: Business Requirements and Critical Success Factors. Prod Oper Manag 29 (4), 935–954.

[32] European Union, 2020. Towards a mandatory EU system of due diligence for supply chains. https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2020)659299.

[33] Frandsen, T., 2017. Evolution of modularity literature: a 25-year bibliometric analysis. International Journal of Operations & Production Management 37 (6), 703–747.

[34] Wang, Y., Ma, H.-S., Yang, J.-H., Wang, K.-S., 2017. Industry 4.0: a way from mass customization to mass personalization production. Advances in Manufacturing 5 (4), 311–320.

[35] Tang, M., Qi, Y., Zhang, M., 2017. Impact of Product Modularity on Mass Customization Capability: An Exploratory Study of Contextual Factors. International Journal of Information Technology & Decision Making 16 (04), 939–959.

[36] Aheleroff, S., Philip, R., Zhong, R.Y., Xu, X., 2019. The Degree of Mass Personalisation under Industry 4.0. Procedia CIRP 81, 1394–1399.

[37] Chouinard, U., Pigosso, D.C., McAloone, T.C., Baron, L., Achiche, S., 2019. Potential of circular economy implementation in the mechatronics industry: An exploratory research. Journal of Cleaner Production 239, 118014.

[38] Arshinder, K., Kanda, A., Deshmukh, S.G., 2011. A Review on Supply Chain Coordination: Coordination Mechanisms, Managing Uncertainty and Research Directions, in: Choi, T.-M., Cheng, T.E. (Eds.), Supply Chain Coordination under Uncertainty. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 39–82.

[39] Cao, M., Zhang, Q., 2011. Supply chain collaboration: Impact on collaborative advantage and firm performance. Journal of Operations Management 29 (3), 163–180.

[40] Leeuw, S. de, Fransoo, J., 2009. Drivers of close supply chain collaboration: one size fits all? International Journal of Operations & Production Management 29 (7), 720–739.

[41] Christopher, M., Holweg, M., Crum, M., Poist, R., 2011. “Supply Chain 2.0”: managing supply chains in the era of turbulence. International Journal of Physical Distribution & Logistics Management 41 (1), 63–82.

[42] Vogelsteller, F., Buterin, V. EIP-20: Token Standard. Ethereum Improvement Proposals. https://eips.ethereum.org/EIPS/eip-20.

[43] Entriken, W., Shirley, D., Evans, J., Sachs, N., 2018. EIP-721: Non-Fungible Token Standard. Ethereum Improvement Proposals. https://eips.ethereum.org/EIPS/eip-721.

Biography

Fabian Dietrich (*1993) is Ph.D. Candidate (Industrial Engineering) at Stellenbosch University, South Africa and works as Research Associate at ESB Business School, Germany. He holds a M.Eng. (Industrial Engineering) from Stellenbosch University and a M.Sc. (Digital Industrial Management and Engineering) from the ESB Business School of Reutlingen University.

Nils Künster (*1995) is a Research Associate at the ESB Business School. He studied Operations Management and holds a M.Sc. from the ESB Business School of Reutlingen University. His research focuses on artificial intelligence in combination with blockchain technology.

Louis Louw (*1974) has been an associate professor at Stellenbosch University since 2018. Previously Prof. Dr. Louis Louw spent twelve years in industry as a business consultant and senior manager, before joining the university in 2014. His research interests are in the field of Digital Operations & Supply Chain Management, and he is a member of the International Association of Learning Factories.

Daniel Palm (*1967) is Professor for Supply Chain Management at the ESB Business School. Univ.Lektor Prof. Dr.techn. Dipl.-Ing. Daniel Palm is head of the teaching and research center for ValueAdded and Logistics Systems at Reutlingen University and the Reutlingen Center Industry 4.0, RZI 4.0 - a cooperation of the Fraunhofer Institutes IPA and IAO as well as Reutlingen University.

3rd Conference on Production Systems and Logistics

A Hybrid Metric For Navigation Of Autonomous Intralogistics Vehicles In Mixed Indoor And Outdoor Operation

Maximilian Zwingel[1], Andreas Blank[1], Peter Schuderer[2], Jörg Franke[1]
[1]*Institute for Factory Automation and Production Systems, FAU Erlangen-Nürnberg, Germany*
[2]*THI Business School, THI Ingolstadt, Germany*

Abstract

While autonomous guided vehicle systems are increasingly used in homogeneous and structured environments, their use in complex and variable scenarios is usually limited. Established algorithms for the navigation of systems use static maps with deterministic metrics, which can only achieve optimal results in clearly defined environments. In dynamic and extensive deployment scenarios, which are also dependent on a large number of influencing parameters, autonomous intralogistics systems cannot yet be deployed dynamically. One example here is mixed transport between buildings under changing weather conditions.

As a solution for dynamic navigation, we propose a hybrid metric in combination with topological maps and cyclic environmental sensing. Based on a quantification of influencing factors on each intralogistics entity, an optimal and dynamic navigation of every system can be performed at any time. The individual components are implemented in the context of an autonomous tow truck system and evaluated in different application scenarios. The results show significant added value in use cases with sudden weather changes and complex route networks.

Keywords

Autonomous Tow Truck Systems; Topological Maps; Hybrid Metric

1. Introduction

The number of use cases for autonomous systems in intralogistics scenarios is continuously growing. Initially limited to warehousing and commissioning, advances in sensor technology and computing hardware increasingly enable systems to be used in unstructured production and logistics environments. The foundation of navigating such environments is a representation of inherent knowledge by static or cyclically updated maps. Within this metric representation, paths for transport entities can be computed and executed according to different specifications for optimal behavior. Commonly, the shortest path is used for navigation of autonomous transport vehicles. [1]

This approach is optimal and often sufficient for use cases with identically qualified transport vehicles and homogeneous operating environments. Within industrial use cases, however, there are often different kinds of transport vehicles as well as dynamically changing working environments. Especially in the case of routing networks with outdoor areas and heterogenous vehicle classes, purely metric considerations of navigation solutions are no longer sufficient to enable optimal navigation. For complex and changing environments, no established algorithm for optimal navigation exists. [2]

DOI: https://doi.org/10.15488/12123

In this work, we propose the use of a hybrid metric combined with a state-of-the-art heuristic for optimal navigation of autonomous transportation vehicles as a function of environmental factors and structural constraints inside intralogistics transportation networks. The proposed solution builds upon a node-based representation of the working environment with inherent information encoding. In conjunction with cyclic monitoring of relevant environmental and weather parameters, this enables continuous evaluation and adjustment of the optimal navigation of autonomous transport systems. By incorporating an open-source framework for route network design, a user-friendly interface for modifying the basic map material and deviating metric weighting becomes possible.

Due to the high degree of interference of parametrization and real-world systems, the described metric is detailed using an autonomous tow truck system in mixed indoor and outdoor operation and the potential in dynamic application areas with changing environmental conditions is shown. Finally, an evaluation and outlook on further relevant improvements is given.

2. Related Work

In general, the navigation of autonomous transport systems is based on a-priori knowledge implemented as maps of the operational environment, whereby established methods can be differentiated into metric and topological approaches [3]. Metric maps represent the operational environment as a true-dimensional relation of granular basic elements, usually area squares of the size of the necessary sensory resolution. Due to the high information density, this form of map representation is suitable for small areas of operation where high accuracy is required in the context of navigation. Topological maps abstract the work area to a network of nodes that represent references within the operational environment. Edges between nodes represent a relation between locations and thus state transitions of the driverless transport vehicle. Due to the applied reduction of information density, topological maps are suitable as a form of representation of large transportation networks as well as the navigation of driverless transportation systems within them. The use of a topological map also enables the efficient mapping of relevant parameters for the operation of driverless systems by attribute assignment to the corresponding nodes and edges. [4, 5]

The use of topological maps has already been used for the mixed indoor and outdoor deployment of autonomous transport vehicles [6] as well as autonomous cars [7]. In the classical approach, nodes represent relevant reference points within the deployment environment. Edges between the nodes form a relation and thus a physical connection between the nodes. The resulting network can be represented as a graph with directed edges, where the metric distance between the connected nodes is usually used as edge weight. It can be seen as an optimization problem for the navigation between two points of the graph using a defined heuristic. A distinction is made between informed and uninformed search methods. [8] Uninformed search methods consider inherent information, such as the smallest sum of previously considered nodes, when evaluating the next node to be considered. Informed search methods also use additional information from a weighted estimate to achieve a targeted solution to the optimization problem. This can represent, for example, the Euclidean distance of the current node to the target node. Exemplary, the heuristic of the A*-Algorithm is defined by

$$f(x) = g(x) + h\ (x) \tag{1}$$

with f(x) being the cost of the current cell, g(x) being the actual cost of the accumulated edges from the start point to the current node and h(x) being the weighted estimate of the current node to the goal. [9]

To obtain an optimal solution to the navigation problem, the heuristic used must be monotonic. This means that the estimated cost must never be overestimated and the triangle inequality holds. [10]

In the case of homogeneous environments, where the metric between waypoints is directly relative to the effort of driving, a topological representation of the map and the application of the classical A* algorithm on the routing graph can be used to optimally implement a navigation of autonomous transport vehicles. In dynamic environments as well as changing environmental influences, a purely metric view of the route network cannot be used for optimal navigation. To extend the A*-algorithm towards dynamic environments and entities with different capabilities, different approaches have been researched.

[11] introduces a time-space network model for navigation of automated guided vehicle fleets. In addition to the pure consideration of the distance traveled for each vehicle, the heuristic is extended to include the state of motion of the vehicles. The metric for optimal navigation is thus extended for an overall optimum by the possible prediction of occupancy states for individual route sections. However, it does not include environmental influences on the overall navigation and only optimizes on its own knowledge of the transport entity.

[12] considers the significance of weather influences on the driving behavior of cars. Based on recorded and predicted weather effects, the edge weights and thus the distance values of a route network are multiplied by a fixed factor. This allows the use of a uniform metric for determining an optimal path. However, decision variants and different capabilities of various vehicles, as they often exist in mixed indoor and outdoor route networks are not considered. The approach is thus not easily transferable towards autonomous transport vehicles in intralogistics.

A parametrical description of autonomous transport vehicles for the adaptation of navigation algorithms is described in [13]. Adaptive navigation is made possible by linking the physical capabilities of the vehicle as well as the kinematics of the vehicle with the achievable velocity in various movement scenarios.

In the field of mixed indoor and outdoor intralogistics, there are no approaches known to the authors that allow adequate consideration of relevant environmental influences on autonomous transport vehicles for the optimal navigation in dynamic mixed indoor and outdoor operational environments. However, such a heuristic is necessary for holistic optimization of transports in case of highly branched route networks with multiple viable options.

3. Definition of a hybrid metric

In [14] and [15], we have shown the influence of different operational environments on the localization capabilities of autonomous transport vehicles as well as a possibility of inherent information coding for describing the operational environment. The definition of a new hybrid metric is necessary to combine both findings in an optimal navigation algorithm, which is updated upon changes in environmental conditions. To ensure an optimal result, the superordinate heuristic must meet the criterion of monotonicity and the triangle inequality while also taking weather influences on the driving behavior of driverless transport vehicles into account.

The driving characteristics of autonomous transport vehicles are primarily defined by their ability to perceive the environment and interpret the corresponding information [16]. Environmental influences affect different sensors in various ways. Optical sensors such as Light Detection and Ranging (LiDAR) or cameras are affected by particles of different sizes inside their corresponding area of observation. In outdoor environments, these can be exemplary be snowflakes, raindrops, fog or dust particles [17]. Passive optical sensors are also dependent on the illumination intensity of the respective detection area. [18]

In addition to the influences on the sensors used and the resulting detection range with corresponding speed and availability restrictions, the physical characteristics of the vehicle are relevant for operation in different operational environments. Information on the maximum speed, restrictions due to floor coverings, maximum gradient values depending on possible over-freezing as well as wind influence on trailers must be considered for optimal navigation.

Since the consideration of relevant factors is not possible in a pure metric, we propose the use of a hybrid metric considering the influence of all mentioned information. For autonomous transport vehicles all impact factors are directly reflected upon the maximum achievable and achieved speed. To include them in the navigation, the cost of nodes inside the A* algorithm can simply be extended to:

$$g_{hybrid}(x, v) = \sum_{i=1}^{n} \frac{x_i}{v_i} \tag{2}$$

$g_{hybrid}(x,v)$ represents the costs of the nodes visited so far in applying the A* algorithm on the route network. It is calculated as sum of the corresponding route elements (real distance) x_i and the possible speed v_i (with $v_i > 0$) for the considered transport vehicle between the nodes K_i and K_{i-1}. While the factor x_i has the same value at any time due to the physical conditions of the area of operation, v_i includes the vehicles dependence on environmental conditions and the used transport systems.

To allow optimal navigation in extensive route networks, an informed search is necessary. The chosen heuristic must also reflect external influences on the autonomous transport vehicle as well as physical capabilities. It must be monotonic as well as compatible with the triangular inequality in order to guarantee an optimal result of the navigation. In complex deployment environments, the use of classical metrics for the cost estimation function, such as the pure Euclidian distance between the currently considered node and the target point, is only suitable to a limited degree. An extended heuristic that satisfies the conditions for optimal navigation on the fastest path is a simple division of the air distance by the maximum speed of the vehicle on the chosen path. However, in order to avoid weighting of each estimated cost with the same maximum feasible speed (which would again result in the output of the shortest and not necessarily fastest path) a decision criterion for a choice of the used speed classes is necessary. Resource-efficiently, this can be achieved by incorporating the number of adjacent nodes and a classification into indoor and outdoor areas. The following formula is used to calculate the cost estimation function:

$$h_{fast} = \frac{(K_{ges} + 1) * h_{A*}}{(K_{indoor} + 1) * v_{indoor,max} + K_{outdoor} * v_{outdoor,situ}} \tag{3}$$

with

$$v_{indoor,max} \geq v_i \; and \; v_{outdoor,situ} \geq v_i$$

K_{ges} is the number of neighboring nodes of the considered element of the graph. h_{A*} represents the classical metric for the estimation function, for example being the Euclidean distance between the element and the target node. K_{indoor} and $K_{outdoor}$ describe the number of adjacent nodes that are categorized as indoor and outdoor nodes, respectively. $V_{indoor,max}$ is the maximum speed that can be achieved in the indoor area. $V_{outdoor,situ}$ is the maximum speed that is allowed by an autonomous transport vehicle in the dynamic outdoor area under the environmental conditions prevailing at the time of calculating the navigation solution. According to the above formula, outdoor nodes that have a similar distance to the destination node are preferentially expanded under favorable weather conditions. The different speeds are again directly dependent on the capabilities of the transport vehicle used as well as the operational environment. The mentioned conditions ensure the monotonicity of the selected metric.

In summary, the heuristic for navigating the fastest path in a given route network with information on current environmental conditions can be written similar to the classical A*-algorithm:

$$f_{hybrid} = g_{hybrid} + h_{fast} \tag{4}$$

4. Implementation of the proposed metric

The implementation of the hybrid metric is heavily dependent on the physical characteristics of each autonomous transport vehicle as well as the environmental conditions. In the following chapter, one implementation is detailed using a driverless tow truck as an example as shown in Figure 1. The corresponding vehicle has a variety of sensors that use different measurement principles to sense the environment and thus enable autonomous operation. The operational environment with dynamic indoor and outdoor areas is provided by a hybrid topological map in .osm data format. Inherent information coding assigns ground conditions, slopes and relevant environmental influences to each node and edge via fixed flags inside the .osm map. The tow truck is equipped with camera-based optical odometry, wheel-based odometry, LiDAR sensors, and GPS and UWB systems. The tow truck can travel at a maximum speed of 10 km/h in outdoor areas. In indoor areas, the speed is limited to 6 km/h due to national safety regulations.

Figure 1: Autonomous tow truck for the exemplary realisation of the hybrid metric

Relevant environmental conditions for the functioning of the autonomous route tow truck are summarized in Table 1. Different influences on the sensors result in a decreasing ability of the tow truck to perceive the environment and thus result in a necessary reduction of the allowed maximum driving speed. The reduction of the maximum speed was determined empirically in tests.

Table 1: External influences on sensor systems and required behavior of the tow truck

Ambient influence	Affected component	Required behavior
Humidity condensing [19]	Camera (passive, active), LiDAR	Driving outdoors not permitted
Wind speed [20]	Stability of towed trailers	>25km/h: driving outdoors not permitted
		[25; 15] km/h: driving speed reduction outdoors to 5 km/h
Illuminance [21]	Camera (passive)	< 10 lx: driving outdoors not permitted
		[10; 200] lx: driving speed reduction outdoors to 5 km/h
Temperature [22]	Traction, Condensate on optical components	< 0 °C: driving speed reduction outdoors to 5 km/h

The navigation of the autonomous tow truck is implemented in the Robot Operating System (ROS) on a control computer running Ubuntu Linux 20.04. The working environment is provided as a hybrid topological map with a corresponding route network. The individual nodes are defined by their position in the Global Positioning System (GPS) coordinate system. Distances between nodes and thus the lengths of the edges result from the Euclidean distance in the north-south, east-west and elevation direction. The relevant environmental conditions in the outdoor area are recorded cyclically every 30 seconds by a weather station and transmitted to the tow truck via LoRa-Wan. The A*-algorithm with the actual heuristics based on the hybrid metric is implemented as a standalone package in ROS. In order to cope with dynamic changes in environmental conditions, an actuality check is made when reaching individual nodes to see if any thresholds of the configuration have been exceeded. If this is the case, a new navigation from the current location to the selected destination is triggered. This whole set up with environmental sensing, reconfiguring, navigation and interfacing with transport orders and the driving hardware is shown in Figure 2.

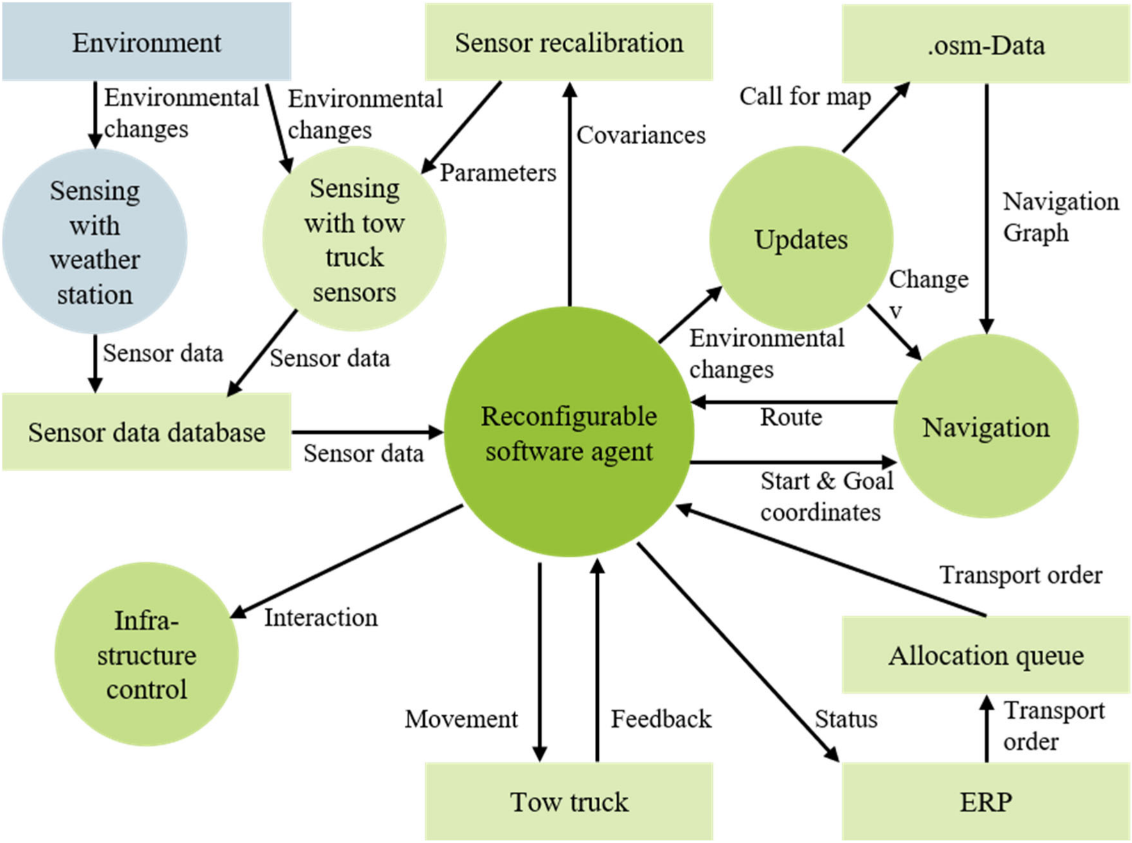


Figure 2: System Setup for the dynamic navigation based on environmental sensing and the hybrid metric

5. Evaluation

To evaluate the functionality of the hybrid metric, its integration described in chapter 4 is tested in different use case scenarios. For the application in dynamic indoor and outdoor areas, the following key features are particularly relevant:

- Correct functionality of the hybrid metric
- Runtime of the algorithm in complex route networks
- Ruggedness of the algorithm for real-time updates

For testing of the functionality of the hybrid metric in rapid changes of environmental conditions, a digital twin of the operational environment and the tow truck was created. Additionally, to the real-world system, this implementation is used to rapidly adjust weather conditions that would be difficult to test in reality. Based on the defined environmental conditions, it was investigated whether the hybrid metric adapts the navigation of the system according to the defined criteria. Figure 3 presents the results of a representative test run. It can be seen that even with sudden weather changes, adjustments to the path are made according to the fastest route.

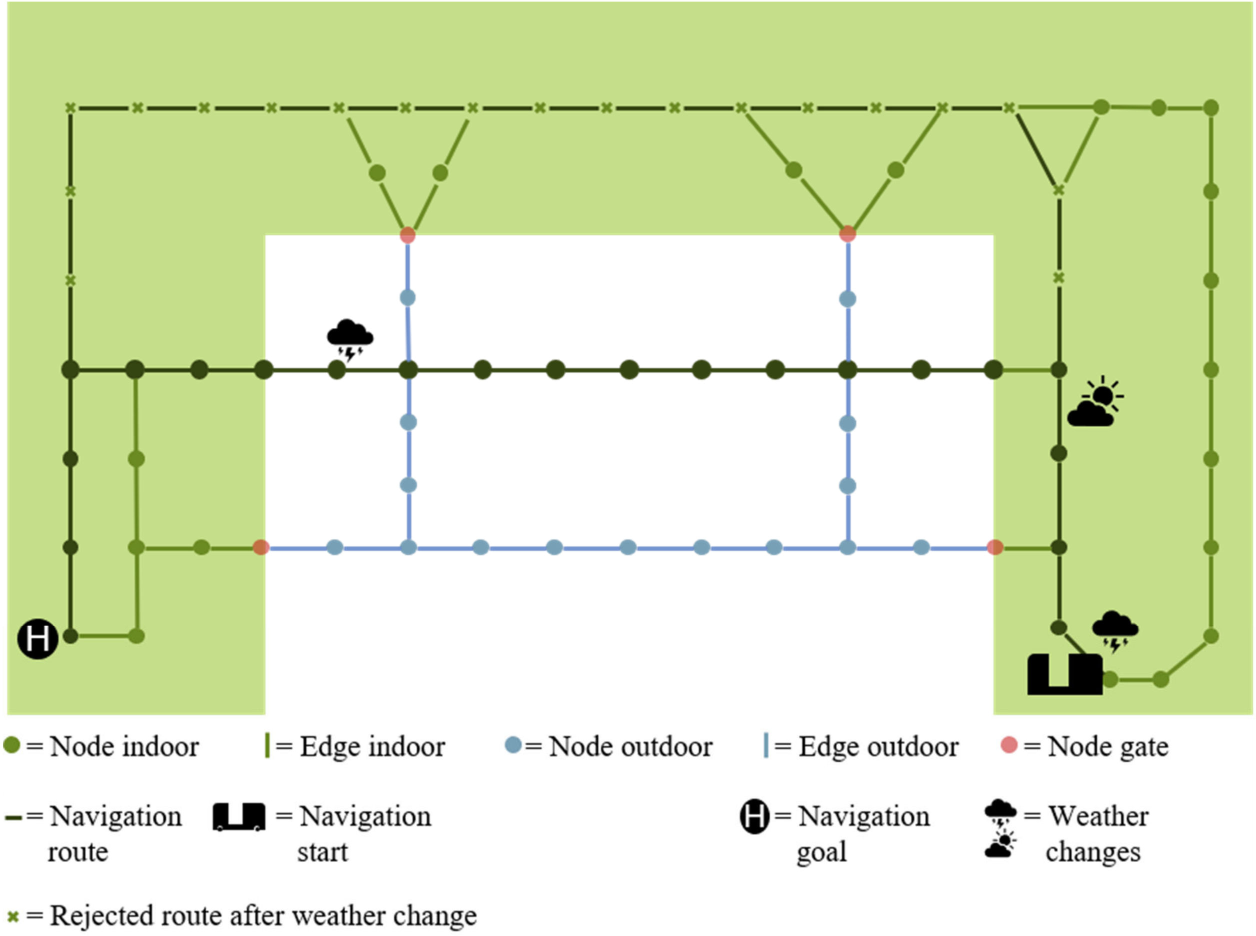


Figure 3: Exemplary visualisation for one evaluation case of the hybrid metric

For an evaluation of the algorithm's runtime in complex route networks, an extensive industrial site with diverse indoor and outdoor areas was digitized and stored in the form of an .osm dataset. On this dataset, different algorithms for randomly generated start and destination points were investigated with respect to their runtime. The proposed heuristic captured the fastest route in every case, while the classical heuristic for an A* algorithm almost exclusively found the shortest route between the two points. The run time of the proposed heuristic is comparable to the application of the classical A* algorithm and scales similar to the values presented in [15].

The robustness of the proposed algorithm was tested during the complete evaluation phase of the research project E|SynchroBot, which results are openly available over the website of the institute FAPS. During the evaluation time, no limitation of the available time of service could be measured.

6. Summary and outlook

In this paper, a hybrid metric for the consideration of dynamic environmental influences on the optimal navigation of autonomous transport vehicles is presented. By considering the specific influences as well as the physical properties of driverless transport entities, the problem of the fastest route of a mixed indoor and outdoor navigation can be solved. The proposed metric is implemented exemplarily on a driverless tow truck showcasing a demonstrative derivation of parameters. In the context of real industrial application scenarios, the qualitative function of the metric could be shown.

In perspective, the use of such a metric allows the autonomous and dynamic navigation of different transport vehicles in extensive areas of application with changing indoor and outdoor areas. A centralized acquisition of environmental data ensures that each vehicle has a corresponding knowledge base for the assessment of external weather influences. Together with the inherent information coding of .osm data, a fast and efficient navigation in the mixed indoor and outdoor logistics context can be achieved.

In perspective, the presented methodologies can be transferred in a variety of use cases. For example, a hybrid consideration of public transport with other means of transport, such as e-scooters or bicycles, would be possible without significant modifications.

As an outlook of the presented work, a further development of the presented heuristic should certainly be mentioned. Due to the monotonicity requirement, a significant underestimation of the actual (time) cost is performed in the current notation. Because this underestimation is applied equally to all nodes, the influences on the order of expansion cancel each other out. However, so far, the result is not suitable for an operational estimation of the remaining transport time. Here, an alternative approach, which makes a real estimate, would be the next step of a fully useful implementation.

Literaturverzeichnis

[1] DUINKERKEN, Mark ; OTTJES, Jaap ; LODEWIJKS, Gabriel: Comparison of Routing Strategies for AGV Systems using Simulation. In: PERRONE, L. Felipe (Hrsg.): *Proceedings of the 2006 Winter Simulation Conference* : *December 3 - 6, 2006, Monterey, CA, U.S.A*. S.l. : Omnipress, 2006, S. 1523–1530

[2] WANG, Chen ; MAO, Jian: Summary of AGV Path Planning. In: *2019 IEEE 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)* : *October 18-20, 2019, Xiamen, China*. Piscataway, NJ : Institute of Electrical and Electronics Engineers Inc, 2019, S. 332–335

[3] KURIC, Ivan ; BULEJ, Vladimir ; SAGA, Milan ; POKORNY, Peter: *Development of simulation software for mobile robot path planning within multilayer map system based on metric and topological maps*. In: *International Journal of Advanced Robotic Systems* 14 (2017), Nr. 6, 172988141774302

[4] JAMALI, Ali ; ABDUL RAHMAN, Alias ; BOGUSLAWSKI, Pawel ; KUMAR, Pankaj ; GOLD, Christopher M.: *An automated 3D modeling of topological indoor navigation network*. In: *GeoJournal* 82 (2017), Nr. 1, S. 157–170

[5] GUO, Lixiao ; YANG, Qiang ; YAN, Wenjun: Intelligent path planning for automated guided vehicles system based on topological map. In: JOELIANTO, Endra (Hrsg.): *2012 IEEE Conference on Control, Systems & Industrial Informatics (ICCSII 2012)* : *Bandung, Indonesia, 23 - 26 September 2012*. Piscataway, NJ : IEEE, 2012, S. 69–74

[6] DARYA, A. M. ; REHMAN, J. ; ALDAWOUD, O. ; SOUDAN, B.: ICAD - An Economical Indoor & Outdoor Delivery Vehicle Capable of Autonomous Navigation. In: *2019 Advances in Science and Engineering Technology International Conferences (ASET)*. Piscataway, NJ : IEEE, 2019, S. 1–6

[7] ORT, Teddy ; PAULL, Liam ; RUS, Daniela: Autonomous Vehicle Navigation in Rural Environments Without Detailed Prior Maps. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)* : IEEE, 2018, S. 2040–2047

[8] ZHANG, Zhanying ; ZHAO, Ziping: *A multiple mobile robots path planning algorithm based on A-star and Dijkstra algorithm*. In: *International Journal of Smart Home* 8 (2014), Nr. 3, S. 75–86

[9] CUI, Shi-Gang ; WANG, Hui ; YANG, Li: A simulation study of A-star algorithm for robot path planning. In: *16th international conference on mechatronics technology*, 2012, S. 506–510

[10] HERNÁNDEZ, Carlos ; MESEGUER, Pedro ; SUN, Xiaoxun ; KOENIG, Sven: Path-adaptive A* for incremental heuristic search in unknown terrain. In: *Nineteenth International Conference on Automated Planning and Scheduling*, 2009

[11] YIN, Shanling ; XIN, Jianbin: Path Planning of Multiple AGVs Using a Time-space Network Model. In: *Proceedings 2019 34rd Youth Academic Annual Conference of Chinese Association of Automation (YAC)* : *June 6-8, 2019, Jinzhou, China*. Piscataway, NJ : IEEE, 2019, S. 73–78

[12] LITZINGER, Paul ; NAVRATIL, Gerhard ; SIVERTUN, Åke ; KNORR, Daniela: Using Weather Information to Improve Route Planning. In: GENSEL, Jérôme (Hrsg.): *Bridging the geographic information sciences* : *International AGILE'2012 conference, Avignon (France), April, 24 - 27, 2012*. Heidelberg : Springer, 2012 (Lecture Notes in Geoinformation and Cartography), S. 199–214

[13] SCHOLZ, Michael ; ZWINGEL, Maximilian ; SCHUDERER, Peter ; FRANKE, Jörg: *Sustainable Intralogistics due to Uniform Software and Modular Transport Entities*. In: *Procedia CIRP* 80 (2019), S. 239–244

[14] HERBERT, M. ; ZWINGEL, M. ; CZAPKA, C. ; FRANKE, J.: A Multi-source Localization System for Driverless Material Transport in Mixed Indoor and Outdoor Areas. In: BEHRENS, Bernd-Arno; BROSIUS, Alexander; DROSSEL, Welf-Guntram; HINTZE, Wolfgang; IHLENFELDT, Steffen; NYHUIS, Peter (Hrsg.): *Production at the Leading Edge of Technology* : *Proceedings of the 11th Congress of the German Academic Association for Production Technology (WGP), Dresden, September 2021*. 1st ed. 2022. Cham : Springer International Publishing; Imprint Springer, 2022 (Springer eBook Collection), S. 421–429

[15] ZWINGEL, Maximilian ; BAIER, Lukas ; BLANK, Andreas ; FRANKE, Jörg; My University (Mitarb.); HERBERGER, David (Mitarb.); HÜBNER, Marco (Mitarb.) : *Hierarchical And Flexible Navigation For AGVs In Autonomous Mixed Indoor And Outdoor Operation*. 2021

[16] YEONG, De Jong ; VELASCO-HERNANDEZ, Gustavo ; BARRY, John ; WALSH, Joseph: *Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review*. In: *Sensors (Basel, Switzerland)* 21 (2021), Nr. 6

[17] BRAHMANAGE, Gayan ; LEUNG, Henry: Outdoor RGB-D Mapping Using Intel-RealSense. In: *2019 IEEE SENSORS*. [Place of publication not identified] : IEEE, 2019, S. 1–4

[18] MÜLLER, Rainer (Hrsg.); FRANKE, Jörg (Hrsg.); HENRICH, Dominik (Hrsg.); KUHLENKÖTTER, Bernd (Hrsg.); RAATZ, Annika (Hrsg.); VERL, Alexander (Hrsg.): *Handbuch Mensch-Roboter-Kollaboration*. München : Hanser, 2019

[19] AKAI, Naoki ; KAKIGI, Yasunari ; YONEYAMA, Shogo ; OZAKI, Koichi: *Development of Autonomous Mobile Robot that Can Navigate in Rainy Situations*. In: *Journal of Robotics and Mechatronics* 28 (2016), Nr. 4, S. 441–450

[20] ZHANG, Qianwen ; SU, Chuqi ; ZHOU, Yi ; ZHANG, Chengcai ; DING, Jiuyang ; WANG, Yiping: *Numerical Investigation on Handling Stability of a Heavy Tractor Semi-Trailer under Crosswind*. In: *Applied Sciences* 10 (2020), Nr. 11, S. 3672

[21] EXPERT, Fabien ; VIOLLET, Stéphane ; RUFFIER, Franck: *Outdoor field performances of insect-based visual motion sensors*. In: *Journal of Field Robotics* 28 (2011), Nr. 4, S. 529–541

[22] MOOREHEAD, Stewart ; SIMMONS, G. R. ; APOSTOLOPOLOUS, D. ; WHITTAKER, William: Autonomous navigation field results of a planetary analog robot in antarctica, Bd. 440. In: *Artificial Intelligence, Robotics and Automation in Space*, 1999, S. 237

Biography

Maximilian Zwingel is a research associate at the Institute for Factory Automation and Production Systems (FAPS) at the FAU Erlangen-Nuremberg with prior research at the TH Ingolstadt since 2018. His research is focussed on autonomous mobile robots in intralogistics environments, especially their sensors and navigation.

Andreas Blank has been a research associate at the Institute for Factory Automation and Production Systems (FAPS) and head of the Industrial and Service Robotics technology field (2014-2021). As PhD student his research is focused on immersive teleoperation, autonomous robot capabilities and human-to-robot skill transfer.

Prof. Dr. Peter Schuderer is a research professor at the University of applied sciences in Ingolstadt (THI). His research focusses on development of innovative methods and techniques to solve business challenges in value-added optimization and flexibilization in production and logistics.

Prof. Dr. Jörg Franke is Head of the Institute for Factory Automation and Production Systems (FAPS) at the FAU. His research focusses on innovative manufacturing processes for mechatronic products. He is involved in leading functions in scientific organisations such as IEEE, CIRP, WGP, and 3D-MID. Before his professorship, he led different management positions in industry e.g. at McKinsey & Co., Robert Bosch GmbH, ZF Lenksysteme GmbH, and Schaeffler AG.

3rd Conference on Production Systems and Logistics

Distributed Manufacturing: A High-Level Node-Based Concept for Open Source Hardware Production

J.C. Mariscal-Melgar[1], Mohammed Omer[1] , Manuel Moritz[1], Pieter Hijma[2], Tobias Redlich[1], and Jens P. Wulfsberg[1]

[1]*Helmut-Schmidt University / University of the Federal Armed Forces, Hamburg, Germany*
[2]*Hamburg Institute for Value Systematics and Knowledge Management, HIWW, Hamburg, Germany*

Abstract

Distributed manufacturing is presented as a means to enable sustainable production and collaboration. Rather than rely on centralised production, distributed manufacturing promises to improve the flexibility and resilience to meet urgent production demands. New frameworks of production, based on manufacturing models with distributed networks, may provide functional examples to industrial practice. This paper discusses efforts in distributed production in the context of Free/Open source hardware and devises a conceptual framework for future pilots at which open source machines, such as a desktop 3D printer, may be manufactured in a network of open/fab lab nodes.

Keywords

Distributed Manufacturing; Microfactory; Open Source Hardware; Free and Open Source; Circular Economy; Open Production

1. Introduction

The study of distributed manufacturing (DM) has become an interesting topic in literature. Among the milieu of topics in this domain of research, this paper is motivated by two interesting developments: First, manufacturing processes are changing over the years [1–3] and so is the environment in which enterprises, social or commercial, operate. Second, the emergence of the Free/Open Source Hardware (OSH) is noteworthy, as it introduces new forms of collaboration in project development, previously limited to the software realm [4]. A third observation is that both these developments have been made possible by a well-established digital infrastructure we now take for granted. The digital infrastructure has allowed Free/Open Source Software to succeed and, in turn, to inspire the OSH movement [5]. Similarly, digital innovations have enabled manufacturing to move beyond centralised conglomerate systems to new forms of distributed networks of production [6] (that we refer to as DM).

These interesting phenomena are topics of discussion in other papers (see Sec. 1.3) but none fully discuss conceptual frameworks that relate DM and OSH and explore the opportunities they can bring to new models of open production that are not limited to 3D printing of parts. In this paper, we provide a conceptual framework for this relationship, motivated by the needs of future pilots of open production. More specifically, we discuss how DM production units can manufacture open source designs in a network of nodes that simplify the high-level DM abstraction, to benefit OSH manufacturing endeavours. To fully understand the relevance of this opportunity, the rest of this section provides background on DM and OSH (Sec. 1.1) and, a discussion on the main advances in manufacturing and global manufacturing issues (Sec. 1.2).

DOI: https://doi.org/10.15488/12171

1.1 OSH and DM: Overview

Collaboration, open knowledge, and open documentation tools in the form of publicly available projects with permissive or copy-left licenses (Free and Open Source licenses [7]) have transformed the social infrastructure in the software realm [8] and may provide the groundwork to re-shape manufacturing networks in the context of OSH. OSH describes machines, devices, tools, and any physical object whose design specifications are made available to end-users and makers so that such physical objects respect the rights of users and allow for collaboration, reuse, modification, and manufacturing of derivative work [9]. Despite the vagueness of the definition and an ongoing standardisation process [10,11] - OSH, in addition to collaborative and open knowledge benefits, may provide economic advantages [12,13] and flexibility in distributed flows of production.

Innovations in infrastructure and production technology have provided manufacturers, small and large, proximity to end-users in distributed capital and material sourcing production networks. DM refers to the nexus of geographically scattered production facilities that are coordinated for the manufacturing of products. Kerdlap et al. [14] define DM as the production, close to points of consumption, of multiple manufacturing sites, both scalable and localised with reduced transport requirements. In our paper, we consider DM as a system of manufacturing in geographically dispersed components divided into sub-parts with production at different geographical distribution networks in a collaborative setting - similar to Rauch et al. [15] but also encompassing the concepts of open production and bottom-up economics [16]. The distribution network may or may not be coordinated by a single entity or technology, and the governance is from one to many stakeholders.

DM depends on modern infrastructure, information, and communication technologies (ICT), advanced enterprise resource planning (ERP), cyber-physical systems (CPS), and internet of things (IoT) in what is described as Industry 4.0 and smart manufacturing trends to integrate supply chains [17]. DM has also evolved from a value-chain of distributed production across various locations to production networks with small, medium, and large collaborating companies with the challenge to operate with the functionality and capacity of highly refined modern conventional mass production sites [18]. Challenges exist in such complex value-chain relationships [19] and the potential of DM in the context of OSH is to provide means to democratise and decentralise manufacturing practices in a global economy - localised and hyper-customised with international networks of collaboration.

1.2 Advances in manufacturing and sustainability issues

Manufacturing has moved beyond traditional processes [3]. Traditional manufacturing is described as one at a time per need basis production tailored to specific individual needs in a barter system. Production method innovations, over centuries, transformed the scale of production and manufacturing landscape from the industrial revolution to standardisation and high-volume mass production - further refined over the years with machine-dominated robotic automation. Manufacturing today operates in complex labour and capital intensive setting with global distribution networks, high-scale manufacturing, and material intensive needs (see [6] for a historical overview).

Global manufacturing networks are not without fault per se but might be taxing when environmental impact, supply shocks, and local economic self-sufficiency are considered. In the context of environmental impact, global manufacturing modern practices have a responsibility to reduce the environmental burden inflicted by processes in product life cycles [20]. Some of the proposed solutions to the sustainability problem suggest enhancing life cycles [21,22], adopting sustainability programs [23,24], or setting up emission frameworks [25,26]. Supply shocks (e.g., financial [27], pandemics [28,29], commodity shocks [30], natural disasters [31], etc.) also exacerbate the risks in manufacturing flow and environmental impact and may provide opportunities for the consideration of sustainability strategies, such as circular economies [29]. As the effects of environmental impact and disruption of the logistic flow of raw/intermediate materials are exacerbated by supply-chain shocks - there is a significant need for an eco-friendly solution to the manufacturing problem that can be geographically dispersed, localised, and conceivably, more adaptable to shocks.

The need for an eco-friendly solution might be tackled by circular economies. A circular economy is described as a closed-loop flow of resources in the production and processing of goods in an environmentally conscious and ideally waste-free manner (see [32,33] for a more exhaustive definition). In the DM context, it means adding circularity to the distributed flow of resources. As society emerges from production and customisation at a large scale [34,35] to personalised production [36,37], distributed manufacturing in a circular economy setting may be one of the sustainable solutions [38].

1.3 Related work

Previous sub-sections provide background information and serve as groundwork to the definition of the role that DM may play in the production and distribution of open-source projects. Works referred to in this section are inspirations to the contents of this paper; as no equivalent idea of a DM nodes network for OSH is available in current literature.

Among related work in the scope of distributed manufacturing, Wittbrodt et al. [39] explore the life cycles of Open Source 3D printers in a distributed manufacturing setting and highlight how 3D printers may make distributed manufacturing feasible. King et al. [40] follow on the idea and discuss open-source technology to overcome production challenges in communities lacking infrastructure and design desktop 3D printers to provide schools and makerspaces with some manufacturing capability. Similarly, Gwamuri et al. [41] discuss a distributed manufacturing model for self-refraction eyeglasses for developing countries and claim the potential to displace centrally manufactured solutions.

Wittbrodt et al. [42] use the term ultra-distributed manufacturing to describe household-scale 3D printing of complex products. The authors study the case of a solar photovoltaic racking system and show how this mode of production could save costs and improve manufacturing quality. Woern et al. [43] assess distributed manufacturing feasibility in the production of flexible products and find economic and technical advantages in using 3D printers in distributed manufacturing.

Redlich et al. [16] explore theoretical underpinnings of bottom-up economics in the context of open production models in co-creation models of production based on collaboration. A concept drawn onto a microfactory model of open production [44], inspired by microfactory experience in the industry [45] (microfactories may be used or not in DM networks, Sec. 2 will explore further). Ellwein et al. [46] analyse distributed manufacturing and identify customisation, cloud manufacturing, digitalisation, and share-economy as factors that bring analogous co-creation collaborations, termed as the separation of design and manufacturing coupled with new ways of cross-border collaboration in distributed re-location of production.

Literature focuses on the theoretical description of complex paradigms and systems of production or the possibilities of manufacturing in household scales. However, there is work to be done on the definition of a framework that may not only be applied to the production of 3D parts but also, the production of machines.

The motivation of this work is to develop a conceptual framework for future pilots of distributed manufacturing and open production models. Hence, this paper proposes a high-level framework for the problem of how production units in DM networks may organise to manufacture open design projects.

The next section will expand on the industrial and community experience when open-source designs meet distributed production - specifically on microfactories and COVID-19 initiatives in distributed manufacturing. This is followed by the definition and description of how decentralised production of OSH machines could look like as a network of fab/open lab nodes.

2. OSH and DM: industrial and community experience

As OSH design files are ideally freely available on the internet for anyone to download, use, modify and repair. OSH may play a vital role in open design distributed manufacturing in both, the end-product and manufacturing equipment [47].

In the case of OSH as end-products, geographically distributed manufacturing, based on these designs, may save financial, human, and time resources required to design and develop a product from scratch [13]. As manufacturers' development costs go down and the nature of OSH allows for re-use and derivative work of products, different manufacturers may not require to re-design from scratch. Collaborative work encourages standardisation and may increase the availability of components. Ideally, this may allow for the cut down of product redundancy, waste, and planned obsolescence, rampant in today's markets [48].

In the case of open-source manufacturing equipment, the same reasoning follows, with the added advantage that local manufacturers in a localised distributed network may source materials, plentiful in supply from shorter-distance logistic connections. The nature of collaborative design in OSH could also differentiate OSH DM from craft production and allow cost-effective localised manufacturing of products.

OSH is not the only innovation needed to realise DM on a localised scale. Multiple, small high-tech factories may ease the challenge that mass production has with large-centralised sites. Microfactories or desktop factories (see [49] for more information on types) stem from their introduction in Japan in the late 1990s as there was a need to cost-effectively produce small precision components on the micro or mini scale without the reliance on large or inflexible manufacturing sites and machine tools [50]. Microfactories challenge traditional manufacturing aimed at economies of scale, mass production, and capital concentration [51]. Microfactories benefit from locality, flexibility, proximity to other sites or customers, lower carbon footprint, hyper-specialisation, and better access and training of skilled labour [45]. Examples of microfactory production networks include desktop factories at Sankyo production of small mechatronic parts [45], Arrival, an electromobility start-up [52], integration of microfactory processes in electric vehicles production [53], and textile industry manufacturing of smart-clothing accessories [54].

Besides microfactory production in a formal production network - small-scale DM was possible during the COVID-19 pandemic. As hospitals faced acute shortages of personal protective equipment (PPE) [55]; innovative community solutions surfaced in localised, small-scale distributed manufacturing efforts. Face-shield jigs were distributed to local communities by coordinating makers to produce batches of laser cut or 3D printed free and open designs [56]. Besides communities, also start-ups were able to pivot their production. An example in Singapore, a small open-source 3D printer farm start-up was able to quickly adapt and supply face-shields to its local neighbourhood [57]. Several other examples exist [58] and illustrate the ability of small-scale distributed production sites to utilise open designs to adapt their production in real-time to satisfy an urgent supply shortage. Flexibility in production provides resilience in the face of disruptions. Hence, traditional mass production capability may be complemented by utilising local 3D printing capacity for the distributive manufacturing of medical equipment [59]. The same case may apply to other industries as OSH and small-scale DM model of production proliferate in the industry.

3. DM of OSH: A conceptual framework example of machine manufacturing

This section will present a conceptual framework based on a theoretical example of how machines may be manufactured in a DM setting. It is assumed that the source of the machine is freely available and that fab/open labs can join production efforts in a locality or logistic network. To simplify the depiction of such a production network – authors use the example of an existing 3D printer from the Fab City Hamburg project. Background to circular economy initiatives will be provided, in particular, the desktop 3D printer from Fab

City Hamburg. This is followed by an applied conceptual framework to the theoretical node DM production network of such a printer.

Fab City is an initiative that began in 2014 in Barcelona with the premise to transform Barcelona into a circular economy – where by 2054, the city should be able to produce everything the city needs locally [61]. Hamburg became the first German city to pledge to become a fab city [62]. Redlich et al. [44] applied concepts of new patterns of value creation (e.g., networking, collaboration, decentralisation, and bottom-up economics) to describe open labs as distributed open-source microfactories able to set up and replicate manufacturing space and resources to participate in a value creation process that is free of hierarchy and regionally or globally connected. With the rise of the maker scene in the hardware space, fabrication laboratories or fab labs have become a new space for participatory digital manufacturing, whereby local makers, students, entrepreneurs, or anyone may join and make almost anything. Fab labs offer a high degree of creativity and project customisation in a friendly space equipped with minimal tools to assist users and makers in development from design and documentation to prototyping and individualised fabrication [44,60].

Open labs are equipped with machine tools or production technologies whose plans, build instructions, bill of materials (BOM), design files and documentation are freely available. As part of the Fab City Hamburg project, the Open Lab Starter Kit (OLSK) is a project that aims to set up a blueprint of free and open-source machines necessary to establish a digital manufacturing Open Lab [63]. The project originates in the vision to create a distributed network of ‘circular’ production in Hamburg. Local fab labs and open labs are equipped with a range of digital manufacturing machines (Laser Cutters, 3D printers, CNC routers) to provide minimal prototyping or micro-factory capability within the establishments. Machines are designed and developed using distributed control versioning tools (e.g., GIT) and hosted in publicly accessible repositories.

The first machine developed is a desktop 3D printer. The design of the printer is based on an open-source design, adapted to fit local design requirements, and made available in a public repository [64]. The developed 3D printer is planned to be distributed to various makerspaces and institutions throughout Hamburg, with the aim of diffusing digital fabrication technologies throughout the city. The Open Lab Starter Kit project aims to produce and distribute other fabrication machines to labs across the city of Hamburg with in-house developed OSH with easily accessible parts and following the requirements of the Open Source Hardware Association (OSHWA) [65]. The OLSK is a suitable platform to devise a framework on how OSH projects could be produced in a localised distributed way. As a minimal example to answer the question of whether an OSH machine can be produced distributively, the authors consider the case of a simple desktop 3D printer.

Several components of the OLSK 3D printer are 3D printed and metallic raw materials are locally sourced and machined. For instance, the base of the printer is machinable with a laser cutter, whereas, the rest of the printer requires conventional hand tools. Hence, makerspaces, with the minimal set of digital manufacturing tools, may be able to replicate machines. Table 1 summarises the different sub-assemblies required by the 3D printer.

Table 1: OLSK 3D printer manufacturing process difficulty by sub-assembly

Sub-Assembly	**Manufacturing process**	**Difficulty**
3D printed parts	3D printing	Easy
Frame and structure	Cutting and drilling	Easy

Sub-Assembly	**Manufacturing process**	**Difficulty**
Enclosure panels	Laser cutting	Moderate
Electronics and wiring	Soldering and crimping	Moderate/Difficult
Testing	Test station	Moderate
Final assembly	Assembly	Moderate
Software	Flashing firmware	Easy

OLSK 3D printers were built in workshops in Hamburg. It was noted that while manufacturing the 3D printers in our labs, component quality was dependent on the skill set of technicians, troubleshooting ability, the type of machines, features, calibration, and quality of the documentation. In terms of the process from prototyping to manufacturing (see [66] for an overview of production planning topologies), figure 1 summarises the activities that could be equally adaptable to other OSH projects.

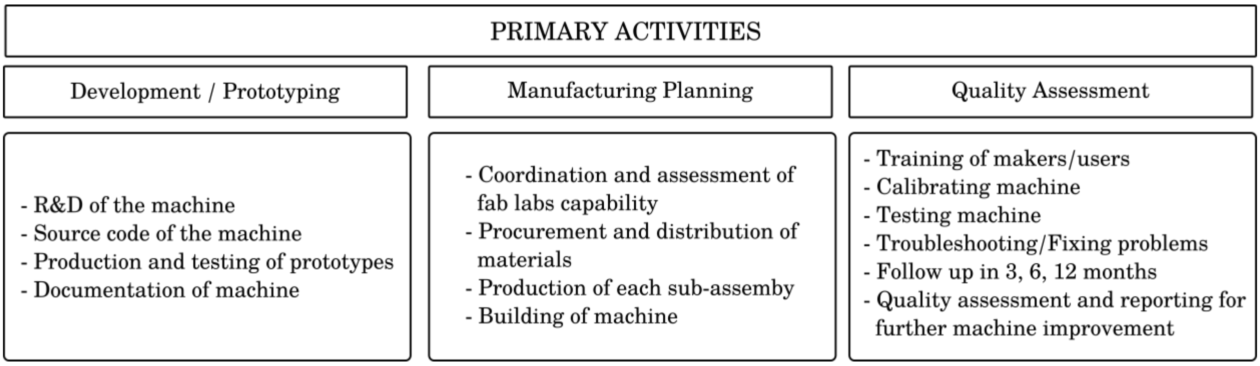


Figure 1: Prototyping, manufacturing and quality control activities

As part of a future pilot of DM production utilising the capacity of localised open labs - the DM of machines may use a network of interconnected nodes providing capacity, capability, flexibility, and high

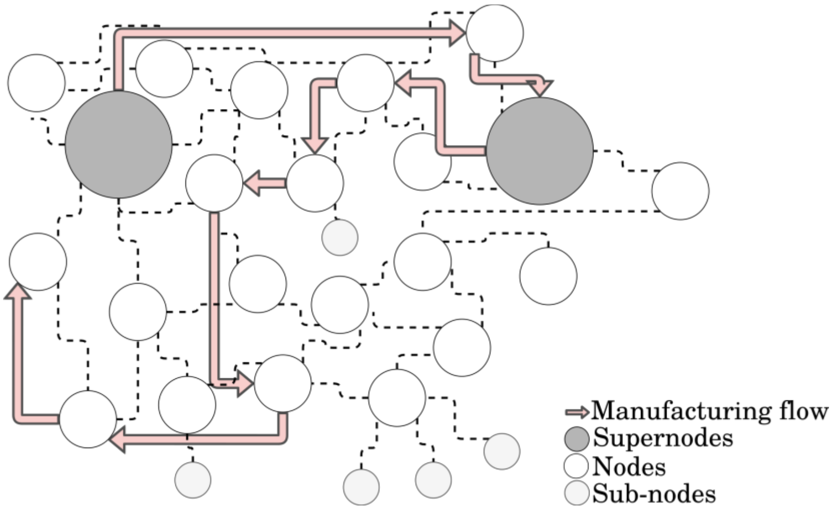


Figure 2: Generalised DM OSH nodes network

customisation/specialisation of manufacturing steps. Figure 2 illustrates the concepts how OLSK machines could be manufactured locally by utilising existing microfactory capability within fab/open labs in the city.

We consider a node as a microfactory in the form of a fab or open lab, able to produce, test, and manufacture a part, component, or sub-assembly of the machine. Nodes are distributed across a city and can undertake one or more work packages as required. The work across nodes might be distributed according to the product to maximise efficiency, minimise logistic travel and reduce waste.

A node is a simple representation of a microfactory that encompasses and simplifies other external factors, inherent to the running of such a micro-production unit (e.g., employees, marketing/business plans, monetary flows). Hence, a node does not address such issues, but it represents on a higher level, no matter the internal configuration or running of it, a microfactory able to produce, test, and manufacture parts assigned to it.

Nodes in the DM OSH production network may act as manufacturers of new nodes if they have the capability of such. Nodes may collaborate in the production, assembly, and flow of materials, parts, sub-assemblies, and products. A system with small-scale production is flexible and resilient against supply-shock production flows. The manufacturing of one machine, for example, may be divided between nodes in the network so that each node produces a part or sub-assembly according to the capabilities and capacity requirements. A node may become a sub-node, which are nodes capable of taking over the production task of another node. Nodes may or may not be part of the DM network of a specific machine. This means that nodes networks may participate in the production of more than one machine and may not be restricted to the production flow of one manufacturing project. Similarly, there may be nodes capable of producing all the manufacturing steps for the machine at once, these are super-nodes and may or may not participate in one or more sub-assemblies of a machine.

The flow of materials and parts may be traceable and trackable through digital product passports. Digital product passports may provide an intelligent, verifiable way to support such relationships (this is depicted by the traced lines connecting the nodes). However, challenges in this field remain. There are stakeholder policy issues [67], requirement composition challenges from an enterprise perspective [68], digital product passport policy guidance such as the upcoming EU Battery Regulation [69], implementation initiatives such as in Hamburg [70], Prague [71], Glasgow [72] or Brussels [72] (for an exhaustive list see the EU Joint Research Centre (JRC) [73]).

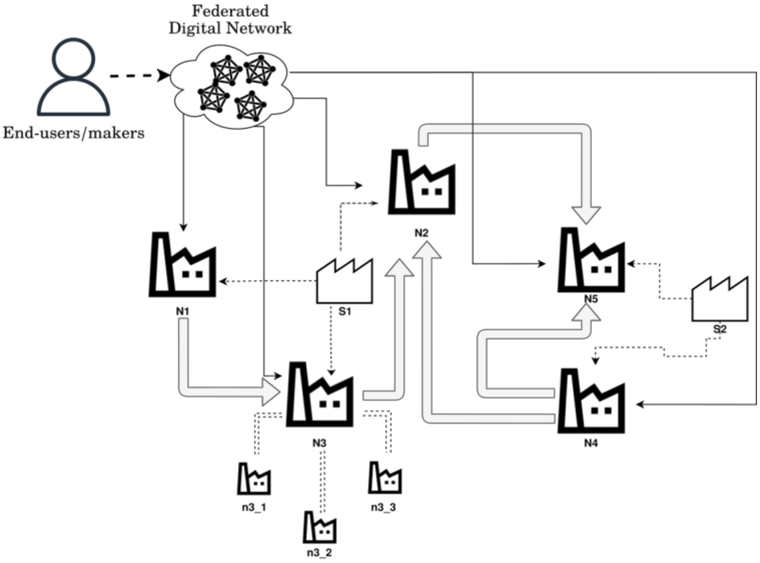


Figure 3: OLSK DM network representation

Such a complex network of nodes may also require high optimisation capability to be able to assess the most optimal way to plan production, a topic of current research, and with work in progress algorithms [74] [75].

A simplified representation of how such a network of nodes may be applied in a pilot is illustrated in Figure 3, where:

- *N1*: Frame and structure producing node
- *N2*: Enclosure panels and pre-assembly node
- *N3*: 3D printing node
- *n3_1*, *n3_2*, *n3_3*: 3D printing sub-nodes
- *N4*: Electronics and wiring node
- *N5*: Final assembly, electronic testing, and software flashing node
- *S1*, *S2*: Local part suppliers

Nodes *N1*, *N2*, *N3*, *N4,* and *N5* exchange information by using the digital infrastructure and produce each sub-assembly or step independently of each other. Production may flow linearly 1-1 as illustrated from *N1* to *N3* or with multiple inter-dependencies as in *N4* to *N2* and *N5*. As expected, the system provides some resilience as sub-nodes *n3_1*, *n3_2*, *n3_3* may backup *N3* if required. As the network grows, more sub-nodes may replace other microfactory nodes.

The cloud in Figure 3 represents a federated digital network, a network infrastructure that is distributed and decentralised across instances of a service, in this case, a DM network. The role of the federated network is to provide the tools to track and trace production flow, decentralise, encourage open participation, and provide means of good exchange supporting both, barter and traditional accounting-based systems (see [76] for more details on resource, event, and agent models). This may provide means of recycling, re-use, and minimising waste in a circular economy model. Additionally, the digital network may be complemented by web3.0 technology: blockchain, decentralised autonomous organisations, decentralised finance, self-sovereignty, and privacy tools [77]. However, as technology evolves, it is unclear how such a system and governance may look in practice on a regional level.

Distribution networks, logistics, e-commerce platforms, and stakeholder relationships, due to simplicity of exposition, are not included in Figure 3 but the manufacturing flows may theoretically be adaptable to changes in such factors. Similarly, nodes with non-manufacturing tasks may also be included in the future (i.e., small-scale specialised transportation distribution nodes or federated instances of e-commerce nodes). As with regards to suppliers and proximity mapping for larger-scale production networks, models such as know-how proximity data matrices [78] may be suitable for waste reduction and sustainability applications. However, further research in this field, in the context of DM, is needed.

DM of OSH may assume the node configuration described in this section. Node networks may grow separately from each other in different regions or grow organically in one region towards other regions. The distributed nature of this system allows future communication between geographically distant networks of nodes. Hence, any fab or open lab in such a network can participate in the manufacturing process. As networks of nodes expand, they are not confined to limited regions of production and may communicate cross-regionally or globally with each other if required. The node configuration provides a high-level proposition on how regions may organise their fab/open labs in a democratised and open production flow. Democratised as producers may have access to such network of production. Open production as nodes may participate in the manufacturing of any OSH part, sub-assembly, or project they may be capable to produce.

As DM enables democratisation [79] and the industry continues to innovate global production networks and value creation - OSH may play a vital role. Open source machines, like the OLSK, in the mesh of fab/open labs, are a starting ground for future endeavours in building distributed manufacturing networks. In the future, local/regional value creation systems may have the ability to become circular economies, self-governing, self-reliant, and resilient to supply shocks. Fab City movement may foresee this trend emerging at the city level, but the nature of open distributed networks is not limited to hard-bound perimeters, it is global.

The suggestions in this paper support effort in localising production in a distributed manner and encourages research-led endeavours to explore alternative production networks that aim to be open, sustainable, and adaptable. The authors propose a node-organisation framework of fab/open labs able to manufacture a desktop 3D printer and likewise, other machines by that same approach. The extent to which such possibilities might or might not apply in practice is thought-provoking, particularly, when compared to existing manufacturing literature and current capacity/capabilities in the global economy. Further work is needed in the form of pilots and case studies to evaluate to what extent is distributed manufacturing of open source hardware ready to undertake such an open production endeavour.

4. Further research

Localised distributed manufacturing in open production spaces is referred to as a new category of production [6]. Distributed manufacturing models in open production spaces lack production planning models [80] and there is a need for research on open distributed production algorithms. Hence, future research may expand on how the conceptual framework, presented in this paper, will organise each of the production node's scheduling, capacity allocation, and logistic provisioning.

The distributed manufacturing of open-source machines, such as our desktop 3D printer, utilising the machine capacities of the geographically distributed makerspaces within a city would also serve as an ideal pilot project to test and evaluate production strategies for optimum utilisation of the production capability of networked open production spaces.

Production allocation and fab/open labs manufacturing specialisations may be of interest for future research too. For instance, is it sensible to divide a manufacturing task into sub-assemblies and distribute sub-assembly production to different microfactory nodes? Or is it more sensible to allow each microfactory to become a super-node and produce one machine each? Furthermore, questions remain about the role of ICT implementations of inter-connected systems making use of web3.0 and Industry 4.0 technologies. Similarly, the role of policymakers and green agendas. There will be trade-offs between manufacturing capacity, efficiency, capability, and sustainability. All these issues add to the role of OSH licensing and certifications, which may be a major setback to real-life implementations of such policies and projects.

Acknowledgements

The authors would like to thank dtec.bw – Digitalization and Technology Research Center of the Bundeswehr, FabCity project, for their support.

The authors also thank the INTERFACER Hamburg project for their support.

References

[1] Browne J, Sackett PJ, Wortmann JC. Future manufacturing systems—Towards the extended enterprise. Computers in Industry 1995;25:235–254. https://doi.org/10.1016/0166-3615(94)00035-O.

[2] Esmaeilian B, Behdad S, Wang B. The evolution and future of manufacturing: A review. Journal of Manufacturing Systems 2016;39:79–100. https://doi.org/10.1016/j.jmsy.2016.03.001.

[3] Wang B. The Future of Manufacturing: A New Perspective. Engineering 2018;4:722–728. https://doi.org/10.1016/j.eng.2018.07.020.

[4] Serrano J. Open Hardware and Collaboration. Proceedings of the 11st Int Workshop on Personal Computers and Particle Accelerator Controls 2017;PCaPAC2016:6 pages, 0.114 MB. https://doi.org/10.18429/JACOW-PCAPAC2016-THKTPLK01.

[5] Gibb A. Building open source hardware: DIY manufacturing for hackers and makers. Pearson Education; 2014.

[6] DeVor RE, Kapoor SG, Cao J, Ehmann KF. Transforming the Landscape of Manufacturing: Distributed Manufacturing Based on Desktop Manufacturing (DM)2. Journal of Manufacturing Science and Engineering 2012;134. https://doi.org/10.1115/1.4006095.

[7] Carver BW. Share and Share Alike: Understanding and Enforcing Open Source and Free Software Licenses. Berkeley Technology Law Journal 2005;20:443. https://heinonline.org/HOL/Page?handle=hein.journals/berktech20&id=469&div=&collection=.

[8] Ueda M. Licenses of Open Source Software and their Economic Values. In: . 2005 Symposium on Applications and the Internet Workshops (SAINT 2005 Workshops), 2005, p. 381–3. https://doi.org/10.1109/SAINTW.2005.1620054.

[9] Moritz M, Redlich T, Grames PP, Wulfsberg JP. Value creation in open-source hardware communities: Case study of Open Source Ecology. In: . 2016 Portland International Conference on Management of Engineering and Technology (PICMET), 2016, p. 2368–75. https://doi.org/10.1109/PICMET.2016.7806517.

[10] Bonvoisin J, Mies R, Boujut J-F, Stark R. What is the “Source” of Open Source Hardware? Journal of Open Hardware 2017;1.

[11] Arndt F, Bonvoisin J, Burkert T, Schattenhofer L, Vos Jd, Flüchter F, et al. DIN SPEC 3105 2021. https://doi.org/https://dx.doi.org/10.31030/3173063.

[12] Pearce JM. Economic savings for scientific free and open source technology: A review. Hardwarex 2020;8:e00139. https://doi.org/10.1016/j.ohx.2020.e00139.

[13] Moritz M, Redlich T, Günyar S, Winter L, Wulfsberg JP. On the Economic Value of Open Source Hardware – Case Study of an Open Source Magnetic Resonance Imaging Scanner. Journal of Open Hardware 2019;3:2. https://doi.org/10.5334/joh.14.

[14] Kerdlap P, Purnama AR, Low JSC, Tan DZL, Barlow CY, Ramakrishna S. Comparing the environmental performance of distributed versus centralized plastic recycling systems: Applying hybrid simulation modeling to life cycle assessment. Journal of Industrial Ecology 2021;n/a. https://doi.org/10.1111/jiec.13151.

[15] Rauch E, Dallinger M, Dallasega P, Matt DT. Sustainability in Manufacturing through Distributed Manufacturing Systems (DMS). Procedia Cirp 2015;29:544–549. https://doi.org/10.1016/j.procir.2015.01.069.

[16] Redlich T, Moritz M. Bottom-up Economics. Foundations of a Theory of Distributed and Open Value Creation. In: Ferdinand J-P, Petschow U, Dickel S, editors. The Decentralized and Networked Future of Value Creation: 3D Printing and its Implications for Society, Industry, and Sustainable Development, Cham: Springer International Publishing; 2016, p. 27–57. https://doi.org/10.1007/978-3-319-31686-4_3.

[17] Pipan M, Protner J, Herakovič N. Integration of Distributed Manufacturing Nodes in Smart Factory. In: Borangiu T, Trentesaux D, Thomas A, Cavalieri S, editors. Service Orientation in Holonic and Multi-Agent Manufacturing, Cham: Springer International Publishing; 2019, p. 424–35. https://doi.org/10.1007/978-3-030-03003-2_33.

[18] Srai JS, Kumar M, Graham G, Phillips W, Tooze J, Ford S, et al. Distributed manufacturing: scope, challenges and opportunities. International Journal of Production Research 2016;54:6917–6935. https://doi.org/10.1080/00207543.2016.1192302.

[19] Krenz P, Stoltenberg L, Markert J, Saubke D, Redlich T. The Phenomenon of Local Manufacturing: An Attempt at a Differentiation of Distributed, Re-distributed and Urban Manufacturing. In: Andersen A-L, Andersen R, Brunoe TD, Larsen MSS, Nielsen K, Napoleone A, et al., editors. Towards Sustainable Customization: Bridging Smart Products and Manufacturing Systems, Cham: Springer International Publishing; 2022, p. 1014–22.

[20] O'Brien C. Global manufacturing and the sustainable economy. International Journal of Production Research 2002;40:3867–3877. https://doi.org/10.1080/00207540210157169.

[21] Mengarelli M, Marconi M, Germani M. A Lifecycle Enhanced Global Manufacturing Platform for Enterprises. Procedia Cirp 2016;52:192–197. https://doi.org/10.1016/j.procir.2016.07.022.

[22] Kara S, Manmek S, Herrmann C. Global manufacturing and the embodied energy of products. Cirp Annals 2010;59:29–32. https://doi.org/10.1016/j.cirp.2010.03.004.

[23] Chun Y, Bidanda B. Sustainable manufacturing and the role of the International Journal of Production Research. International Journal of Production Research 2013;51:7448–7455. https://doi.org/10.1080/00207543.2012.762135.

[24] Golini R, Longoni A, Cagliano R. Developing sustainability in global manufacturing networks: The role of site competence on sustainability performance. International Journal of Production Economics 2014;147:448–459. https://doi.org/10.1016/j.ijpe.2013.06.010.

[25] Gurtu A, Searcy C, Jaber MY. A Framework for Reducing Global Manufacturing Emissions. The Journal of Environment & Development 2016;25:159–190. https://doi.org/10.1177/1070496515623821.

[26] Chryssolouris G, Papakostas N, Mavrikios D. A perspective on manufacturing strategy: Produce more with less. Cirp Journal of Manufacturing Science and Technology 2008;1:45–52. https://doi.org/10.1016/j.cirpj.2008.06.008.

[27] Baldwin R. The Greater Trade Collapse of 2020: Learnings from the 2008-09 Great Trade Collapse. VoxeuOrg 2020. https://voxeu.org/article/greater-trade-collapse-2020 (accessed December 28, 2021).

[28] Barua S. Understanding Coronanomics: The Economic Implications of the Coronavirus (COVID-19) Pandemic. Rochester, NY: Social Science Research Network; 2020. https://doi.org/10.2139/ssrn.3566477.

[29] Ibn-Mohammed T, Mustapha K, Godsell J, Adamu Z, Babatunde K, Akintade D, et al. A critical analysis of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies. Resources, Conservation and Recycling 2021;164:105169. https://doi.org/10.1016/j.resconrec.2020.105169.

[30] Krokida S-I, Lambertides N, Savva CS, Tsouknidis DA. The effects of oil price shocks on the prices of EU emission trading system and European stock returns. The European Journal of Finance 2020;26:1–13. https://doi.org/10.1080/1351847X.2019.1637358.

[31] Park Y, Hong P, Roh JJ. Supply chain lessons from the catastrophic natural disaster in Japan. Business Horizons 2013;56:75–85. https://doi.org/10.1016/j.bushor.2012.09.008.

[32] Geissdoerfer M, Savaget P, Bocken N, Hultink EJ. The Circular Economy - A New Sustainability Paradigm? Rochester, NY: Social Science Research Network; 2017. https://papers.ssrn.com/abstract=2930842 (accessed January 3, 2022).

[33] Kirchherr J, Reike D, Hekkert M. Conceptualizing the circular economy: An analysis of 114 definitions. Resources, Conservation and Recycling 2017;127:221–232. https://doi.org/10.1016/j.resconrec.2017.09.005.

[34] Kotha S. Mass customization: Implementing the emerging paradigm for competitive advantage. Strategic Management Journal 1995;16:21–42. https://doi.org/10.1002/smj.4250160916.

[35] Alm R, Cox WM. The right stuff: America's move to mass customization. Annual Report 1998:3–26. https://econpapers.repec.org/article/fipfeddar/y_3a1998_3ap_3a3-26.htm (accessed December 28, 2021).

[36] Wang Y, Ma H-S, Yang J-H, Wang K-S. Industry 4.0: a way from mass customization to mass personalization production. Advances in Manufacturing 2017;5:311–320. https://doi.org/10.1007/s40436-017-0204-7.

[37] Aheleroff S, Mostashiri N, Xu X, Zhong RY. Mass Personalisation as a Service in Industry 4.0: A Resilient Response Case Study. Advanced Engineering Informatics 2021;50:101438. https://doi.org/10.1016/j.aei.2021.101438.

[38] Turner C, Moreno M, Mondini L, Salonitis K, Charnley F, Tiwari A, et al. Sustainable Production in a Circular Economy: A Business Model for Re-Distributed Manufacturing. Sustainability 2019;11:4291. https://doi.org/10.3390/su11164291.

[39] Wittbrodt BT, Glover AG, Laureto J, Anzalone GC, Oppliger D, Irwin JL, et al. Life-cycle economic analysis of distributed manufacturing with open-source 3-D printers. Mechatronics 2013;23:713–726. https://doi.org/10.1016/j.mechatronics.2013.06.002.

[40] King DL, Babasola A, Rozario J, Pearce JM. Mobile Open-Source Solar-Powered 3-D Printers for Distributed Manufacturing in Off-Grid Communities. Challenges in Sustainability 2014;2:18–27. https://doi.org/10.12924/cis2014.02010018.

[41] Gwamuri J, Wittbrodt BT, Anzalone NC, Pearce JM. Reversing the Trend of Large Scale and Centralization in Manufacturing: The Case of Distributed Manufacturing of Customizable 3-D-Printable Self-Adjustable Glasses. Rochester, NY: Social Science Research Network; 2014. https://papers.ssrn.com/abstract=3330068 (accessed January 25, 2022).

[42] Wittbrodt B, Laureto J, Tymrak B, Pearce JM. Distributed manufacturing with 3-D printing: a case study of recreational vehicle solar photovoltaic mounting systems. Journal of Frugal Innovation 2015;1:1. https://doi.org/10.1186/s40669-014-0001-z.

[43] Woern AL, Pearce JM. Distributed Manufacturing of Flexible Products: Technical Feasibility and Economic Viability. Technologies 2017;5:71. https://doi.org/10.3390/technologies5040071.

[44] Redlich T, Buxbaum-Conradi S, Basmer-Birkenfeld S-V, Moritz M, Krenz P, Osunyomi BD, et al. OpenLabs – Open Source Microfactories Enhancing the FabLab Idea. In: . 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA: IEEE; 2016, p. 707–15. https://doi.org/10.1109/HICSS.2016.93.

[45] Okazaki Y, Mishima N, , Ashida K. Microfactory—Concept, History, and Developments. Journal of Manufacturing Science and Engineering 2005;126:837–844. https://doi.org/10.1115/1.1823491.

[46] Ellwein C, Schmidt A, Lechler A, Riedel O. Distributed Manufacturing: A Vision about Shareconomy in the Manufacturing Industry. In: . Proceedings of the 2019 3rd International Conference on Automation, Control and Robots, New York, NY, USA: Association for Computing Machinery; 2019, p. 90–5. https://doi.org/10.1145/3365265.3365270.

[47] Lowe AS. Distributed Manufacturing: Make Things Where You Need Them. In: Redlich T, Moritz M, Wulfsberg JP, editors. Co-Creation, Cham: Springer International Publishing; 2019, p. 37–50. https://doi.org/10.1007/978-3-319-97788-1_4.

[48] Bonvoisin J. Implications of Open Source Design for Sustainability. In: Setchi R, Howlett RJ, Liu Y, Theobald P, editors. Sustainable Design and Manufacturing 2016, Cham: Springer International Publishing; 2016, p. 49–59. https://doi.org/10.1007/978-3-319-32098-4_5.

[49] Montes JO, Olleros FX. Local on-demand fabrication: microfactories and online manufacturing platforms. Journal of Manufacturing Technology Management 2020;32:20–41. https://doi.org/10.1108/JMTM-07-2019-0251.

[50] Kawahara N, Suto T, Hirano T, Ishikawa Y, Kitahara T, Ooyama N, et al. Microfactories; new applications of micromachine technology to the manufacture of small products. Microsystem Technologies 1997;3:37–41. https://doi.org/10.1007/s005420050052.

[51] Järvenpää E, Heikkilä R, Siltala N, Prusi T, Tuokko R. Micro-factories. In: . Micromanufacturing Engineering and Technology, Elsevier; 2015, p. 549–79. https://doi.org/10.1016/B978-0-323-31149-6.00023-2.

[52] Boudette NE. Arrival Developing Electric Vehicles Without Assembly Line - The New York Times. An EV Start-up Backed by Ups Does Away with the Assembly Line 2021. https://www.nytimes.com/2021/04/21/business/arrival-electric-vehicles.html (accessed January 12, 2022).

[53] Stavropoulos P, Papacharalampopoulos A, Michail C, Vassilopoulos V, Alexopoulos K, Perlo P. A two-stage decision support system for manufacturing processes integration in microfactories for electric vehicles. Procedia Manufacturing 2021;54:106–111. https://doi.org/10.1016/j.promfg.2021.07.017.

[54] Lee S, Rho SH, Lee S, Lee J, Lee SW, Lim D, et al. Implementation of an Automated Manufacturing Process for Smart Clothing: The Case Study of a Smart Sports Bra. Processes 2021;9:289. https://doi.org/10.3390/pr9020289.

[55] Chaib F. Shortage of personal protective equipment endangering health workers worldwide. World Health Organization 2020. https://www.who.int/news/item/03-03-2020-shortage-of-personal-protective-equipment-endangering-health-workers-worldwide (accessed January 1, 2021).

[56] Frazer JS. The Role of Distributed Manufacturing and 3D Printing in Development of Personal Protective Equipment Against COVID-19. In: Sandhu K, Singh S, Prakash C, Sharma NR, Subburaj K, editors. Emerging Applications of 3D Printing During CoVID 19 Pandemic, Singapore: Springer; 2022, p. 15–34. https://doi.org/10.1007/978-981-33-6703-6_2.

[57] Lin O. Singaporean youth produce & donate 3D-printed face shields to hospitals from their own printing startup 2020. https://mothership.sg/2020/04/singapore-3d-face-shield-donate/ (accessed January 5, 2022).

[58] Frazer JS, Shard A, Herdman J. Involvement of the open-source community in combating the worldwide COVID-19 pandemic: a review. Journal of Medical Engineering & Technology 2020;44:169–176. https://doi.org/10.1080/03091902.2020.1757772.

[59] Manero A, Smith P, Koontz A, Dombrowski M, Sparkman J, Courbin D, et al. Leveraging 3D Printing Capacity in Times of Crisis: Recommendations for COVID-19 Distributed Manufacturing for Medical Equipment Rapid Response. International Journal of Environmental Research and Public Health 2020;17:4634. https://doi.org/10.3390/ijerph17134634.

[60] Mikhak B, Lyon C, Gorton T, Gershenfeld N, McEnnis C, Taylor J. FAB LAB: AN ALTERNATE MODEL OF ICT FOR DEVELOPMENT 2002:7.

[61] Fab City Initiative. Fab City Global Initiative 2014. https://fab.city/ (accessed January 11, 2022).

[62] Fab City Hamburg. Fab City Hamburg Lokal produziert, global vernetzt. Fab city hamburg 2019. https://www.fabcity.hamburg/ (accessed January 11, 2022).

[63] OLSK H. Open Lab Starter Kit (OLSK) Hamburg. Olsk 2022. https://web.archive.org/web/20220118101124/https://hardware.development.fabcity.hamburg/open-lab-starter-kit/ (accessed January 18, 2022).

[64] Fab City Hamburg. Open Lab Starter Kit. Gitlab 2022. https://gitlab.fabcity.hamburg/hardware/open-lab-starter-kit (accessed January 18, 2022).

[65] OSHWA. OSHWA Definition 2021. https://www.oshwa.org/definition/.

[66] Zijm W. Towards intelligent manufacturing planning and control systems. Or-Spektrum 2000;22:313–345. https://doi.org/10.1007/s002919900032.

[67] Adisorn T, Tholen L, Götz T. Towards a Digital Product Passport Fit for Contributing to a Circular Economy. Energies 2021;14:2289. https://doi.org/10.3390/en14082289.

[68] Donetskaya JV, Gatchin YA. Development of Requirements for The Content of a Digital Passport and Design Solutions. Journal of Physics: Conference Series 2021;1828:012102. https://doi.org/10.1088/1742-6596/1828/1/012102.

[69] Walden J, Steinbrecher A, Marinkovic M. Digital Product Passports as Enabler of the Circular Economy. Chemie Ingenieur Technik 2021;93:1717–1727. https://doi.org/10.1002/cite.202100121.

[70] Fab City Hamburg. INTERFACER Project. Interfacer Project 2021. https://www.fabcity.hamburg/en/fab-city-os-projekt-interfacer/ (accessed January 18, 2022).

[71] Institute Circular Economy Prague. Circular Prague - Insights - Circle Economy. 2019. https://www.circle-economy.com/resources/circular-prague (accessed January 20, 2022).

[72] Kębłowski W, Lambert D, Bassens D. Circular economy and the city: an urban political economy agenda. Culture and Organization 2020;26:142–158. https://doi.org/10.1080/14759551.2020.1718148.

[73] EU JRC. JRC City Science Initiative. Jrc Science Hub Communities - European Commission 2022. https://ec.europa.eu/jrc/communities/en/community/3393/library (accessed January 20, 2022).

[74] Lara CL, Bernal DE, Li C, Grossmann IE. Global optimization algorithm for multi-period design and planning of centralized and distributed manufacturing networks. Computers & Chemical Engineering 2019;127:295–310. https://doi.org/10.1016/j.compchemeng.2019.05.022.

[75] Zhang X, Liu X, Tang S, Królczyk G, Li Z. Solving Scheduling Problem in a Distributed Manufacturing System Using a Discrete Fruit Fly Optimization Algorithm. Energies 2019;12:3260. https://doi.org/10.3390/en12173260.

[76] Schwaiger WSA. The REA Accounting Model: Enhancing Understandability and Applicability. In: Johannesson P, Lee ML, Liddle SW, Opdahl AL, Pastor López Ó, editors. Conceptual Modeling, Cham: Springer International Publishing; 2015, p. 566–73. https://doi.org/10.1007/978-3-319-25264-3_43.

[77] Voshmgir S, Wildenberg M, Rammel C, Novakovic T. Sustainable Development Report: Blockchain, the Web3 & the SDGs 2019. https://epub.wu.ac.at/7453/ (accessed January 14, 2022).

[78] Pachot A, Albouy-Kissi A, Albouy-Kissi B, Chausse F. Decision support system for distributed manufacturing based on input-output analysis and economic complexity. Arxiv:220100694 [Cs] 2021. http://arxiv.org/abs/2201.00694 (accessed January 19, 2022).

[79] Okwudire CE, Madhyastha HV. Distributed manufacturing for and by the masses. Science 2021;372:341–342. https://doi.org/10.1126/science.abg4924.

[80] Hildebrandt L, Redlich T, Wulfsberg JP. Production Planning And Control In Distributed And Networked Open Production Sites – An Integrative Literature Review 2021. https://doi.org/10.15488/11292.

Biography

J.C. Mariscal-Melgar, M.Sc., is a developer and maker with an interest in interdisciplinary research in the field of Free/Open source software and hardware. He supports the research efforts at the Institute of Production Engineering (LaFT) at Helmut-Schmidt-University, Hamburg since 2021.

Mohammed Omer, M.Sc., studied mechanical engineering at the Technical University of Dortmund with a focus on manufacturing technologies. He joined the Institute of Production Engineering (LaFT) at Helmut-Schmidt-University in Hamburg in 2021. His research interest is in open source hardware development in the field of machine tools.

Dr.-Ing. Dipl-Ing. Manuel Moritz, MBA is a research assistant at the Institute of Production Engineering at Helmut-Schmidt University, Hamburg and member of the working group value creation systematics. His research focus is value co-creation in online communities with or without organizations involved.

Dr. Pieter Hijma, is a researcher at the Hamburg Institute for Value Systematics and Knowledge Management, HIWW. His computer science background supports his research goal to create a digital ecosystem based on open source software to support documentation of open source hardware.

Dr.-Ing. Dipl.-Wirt.-Ing. Tobias Redlich, MBA, is the senior engineer at the Institute of Production Engineering (LaFT) in the Department of Mechanical Engineering at Helmut-Schmidt-University in Hamburg since 2013. His research focuses on value creation systematics and the impact of new value creation patterns on production engineering and production management.

Univ.-Prof. Dr.-Ing. Jens Wulfsberg, is head of Chair of Production Engineering in the Department of Mechanical Engineering at Helmut-Schmidt- University in Hamburg since 2001. The main research areas at the Manufacturing Technology Laboratory (LaFT) are value creation systems, robotics, micro manufacturing as well as additive manufacturing and lightweight construction

3rd Conference on Production Systems and Logistics

Explainable Deep Reinforcement Learning For Production Control

Philipp Theumer[1], Florian Edenhofner[1], Roland Zimmermann[1], Alexander Zipfel[1]
[1]*Fraunhofer IGCV, Fraunhofer Institute for Casting, Composite and Processing Technology IGCV, Augsburg, Germany*

Abstract

Due to the growing number of variants and smaller batch sizes manufacturing companies have to cope with increasing material flow complexity. Thus, increasing the difficulty for production planning and control (PPC) to create a feasible and economic production plan. Despite significant advances in PPC research, current PPC systems do not yet sufficiently meet the industry's requirements (e.g., decision quality, reaction time, user trust). However, recent progress in the digitalization of production systems results in an increased amount of data being collected, thus enabling the use of data-intensive applications technologies, e.g., machine learning (ML). ML provides new possibilities for PPC to handle increasing complexity caused by rising numbers of product variants paired with smaller lot sizes. At the same time, ML can increase the decision quality and reduce the reaction time to disturbances in the production system, e.g., machine breakdowns. Partly, ML models, e.g., artificial neural networks (ANN), are perceived as black-box models, resulting in reduced user's trust in the decision proposed by an ML-based PPC system. The approach presented in this publication aims at a more functional and user-friendly PPC system by leveraging multi-agent reinforcement learning (MARL), an accomplished approach within the field of ML-based production control, and approaches for explaining decisions made by reinforcement learning (RL) algorithms. With the help of MARL, short reaction time and high decision quality can be realized. Subsequently, the developed MARL system is combined with methods from the field of explainable Artificial Intelligence (XAI) to increase the users' trust. The use case results show that with the help of the developed system, rule-based controls, which are often used in industry, can be outperformed while providing explainable decisions.

Keywords

Production Control; Machine Learning; Deep Reinforcement Learning; Multi-Agent Reinforcement Learning; Explainable AI

1. Introduction

Manufacturing companies are facing an increasing material flow complexity because of a rising number of variants and a decrease in batch size [1]. More individualized production processes result in a rising production complexity and thus, in challenges (e.g., reaction to disturbances) for efficient production management [2–4]. For example, to determine a cost-optimal sequence, different set-up times, processing times, and necessary process steps for each variant have to be taken into account. With each newly introduced variant, the solution space increases.

In this context, complexity can be differentiated into static and dynamic complexity. While static complexity focuses on the long-term design of production systems, dynamic complexity results from short-term changes in production structures as well as material and information flows triggered by unpredictable disruptions (e.g., machine breakdowns) [2,5].

DOI: https://doi.org/10.15488/12158

The dynamic complexity results in decision-making situations in the context of production control, in which the employee's experience is no longer sufficient to react optimally near real-time while considering the economic effects [6]. With the help of decision support systems, the user can be supported and thus be enabled to react appropriately and avoid adverse effects on the production system like downtime due to a material flow break [6].

Despite these potentials, the decision support of PPC systems is only partially accepted in practice. An indication of this can be seen in the frequency with which manual rescheduling is carried out. In the study conducted by LÖDDING ET AL., only 19 % of respondents stated that the planning of PPC systems is accepted and not overridden [7]. This can be caused by a lack of acceptance of the employees' proposed decisions or a low quality of the systemic proposals due to the production system's high complexity. The lack of trust is also shown in just under 35 % of respondents, who rated their confidence in the PPC system's results [7]. However, KLETTI [6] identifies employee acceptance as crucial for decision support systems in manufacturing. In particular, the user-oriented presentation of information needs to be improved in 71 % of the companies surveyed [8].

In the context of ML, the already existing problem of lack of trust in decision proposals of systems is further aggravated. ML methods, enabled by the increase of data collected, provide new possibilities for PPC to handle the rising complexity. Especially with regard to the dynamic complexity, ML methods allow the increase of decision quality while reducing the reaction time [9]. An accomplished approach for ML-based PPC is using multi-agent systems (MAS) based on deep reinforcement learning (DRL) [10–12]. For this approach, ANNs are used for choosing actions [13]. On the downside, ML models like ANNs are perceived as black-box models [14]. Therefore, the user's trust in decisions proposed by the system can be even further diminished [14].

Within the scope of this paper, an approach for a more functional and user-friendly PPC system is developed. Therefore, a MAS based on DRL is developed to realize short reaction time and high decision quality. For tackling low trust in the proposed decision, approaches from the field of explainable ML are used for realizing the explainability of the system's decisions.

2. State of research

This section focuses on a brief introduction to different approaches leveraging deep reinforcement learning in production control (2.1) and methods for explainable decision finding (2.2).

2.1 Deep reinforcement learning in production control

PPC encompasses an organization's entire materials, time, and production management, as a holistic concept [15]. The target of a PPC system is to increase the logistic performance while maintaining or reducing the logistic costs. High logistic performance is characterized by high delivery reliability and short delivery time. Low inventory and high utilization of machines are influencing the logistic costs beneficially. Due to the competing targets, these must be prioritized on a company-specific basis. PPC play a key role in achieving efficient and economical production [16]. In recent years, approaches leveraging RL have gained much attention within production control, due to the high potential of solving complex problems [10–12].

In RL, an agent interacts with its environment and learns a strategy—also called policy, $\pi(S_t)$—to maximize its reward (R_{t+1}). With the help of the policy, the agent performs an action (A_t) depending on its state (S_t). Based on S_t and A_t the agent receives R_{t+1} [13]. Within the field of RL, deep learning (DL), which uses ANNs, can be used to determine the policy for chosen agents. The combination of RL and DL is called DRL and is a subfield of ML [17]. DRL enables the agents to approximate a function to learn how to behave optimally in an environment and reach the given goals (e.g., high delivery reliability). The agent learns the optimal strategy by choosing actions based on the current state of the environment, with the goal of maximizing a numerical reward [13]. DRL is a promising way to solve problems, which cannot or only with much effort be solved analytically [18].

A MAS consists of a set of agents interacting with the environment to perform one or more tasks jointly. The agents need information about their respective environment to achieve this optimization goal. The agents must obtain information from the environment, evaluate it with respect to the goals, and then select suitable actions [19]. WASCHNECK ET AL. [12] combined a MAS with DRL, realizing a decentralized autonomous approach for a dispatching heuristic, which was successfully tested for a production system producing semiconductors. The agents choose the best possible actions by using a policy based on a Deep Q-Network (DQN) [12]. DQN is based on Q-learning, but an ANN approximates the Q-values instead of the Q-table [20]. The risk of local optimization of the MAS was reduced by using a global reward function [12]. RÖSCH ET AL. [21] implemented a MAS using proximal policy optimization (PPO), a DRL approach, for energy-oriented production control. Agents who had the ability to control the electricity level had to cooperate with electricity-consuming agents to maximize a common reward. Therefore, jobs had to be scheduled to be completed within a given period of time while avoiding a violation of a given energy threshold.

The approaches presented were able to increase decision quality—represented by the improvements of production control—while reducing the reaction time significantly. Neither of the approaches was focusing the explainability of the decisions made by the MAS.

2.2 Explainable AI

Different approaches can realize the explainability of decisions in the context of ML. On the one hand, transparent models can be used; e.g., the parameters used for classification can be read out directly in a decision tree. Thus, the entire decision process is comprehensible, and the model does not require further processing [22]. These models are also called ante-hoc models [23]. However, ante-hoc models are disadvantageous in terms of the model's accuracy compared to opaque models (e.g., ANN) [22].

On the other hand, post-hoc methods can be used to subsequently explain decisions made by opaque models[22,23]. With the help of the post-hoc methods, the decisions of ML algorithms can be explained while levering the advantages concerning the model's accuracy [24]. There are two categories of explainability. Firstly, global explainability describes relationships learned by the model and its general behavior. Secondly, local explainability determines the influences of specific features leading towards a specific prediction [22]. Frequently used post-hoc methods include Local Interpretable Model-agnostic Explanations (LIME) [25] and SHapley Additive exPlanations (SHAP). SHAP is an approach based on the Shapley Value [26] and decomposes a model's prediction into each attribute's contribution to that prediction [27].

REHSE ET AL. [28] use an approach of explainable ML in the context of a model fabric. For this approach, a recurrent neural network is used to make predictions about the further course of the production processes. These predictions are subsequently explained by using LIME. Here, both local (individual decisions) and global explainability (general model's behavior) are realized and subsequently visualized to the user [28]. KUHNLE ET AL. [29] investigate the decision logic of a single agent with DRL using a decision tree. However, the comprehensive explainability of the agents' decisions needs to be further analyzed to overcome the black-box problem [29]. Local and global explainability was used by HUBER ET AL. [30] to explain the behavior of a DQN-agent in a single agent environment. It was investigated that the combination of a local and global explanation helped to achieve a higher performance in the tests conducted [30]. Thus, the combination of MAS with DRL and explainable Artificial Intelligence (XAI) within the production control field promises great potential for improved performance and user-oriented information visualization.

3. Approach

The approach presented within this paper—based on a MAS with DRL— intends to realize short reaction time and high decision quality without neglecting the user's trust in decisions proposed by an agent-based decision support system. Therefore, ANNs are being used to select the best possible action based on the system's current status. Due to the black-box nature of ANNs, additional steps are needed to realize the explainability of decisions. The developed approach is subdivided into two phases. (1) A user-specific observation and action space is defined based on specific user roles within a production system. Subsequently, the multi-agent framework is developed, which enables choosing the best possible action based on the system's current status. (2) Lastly, a method for the explainability of decisions made by the agent and the specific ANN is implemented. Thus, the developed system is enabled to propose explanations.

3.1 Multi-agent system with DRL

This section defines the observation space as well as the action space, which are indispensable for the use of a MAS. The observation space determines the data, which each agent receives for selecting an available action from the action space. For determining the action space for each agent, potential actions (e.g., selection of the following order, short-term capacity increase) are identified for individual user groups in PPC (e.g., production controllers, foremen). Thus, representing their ability to influence the material flow within their user-specific scope in the context of production control. To define these company-specific measures, expert interviews have to be conducted. By defining these possible actions, the decision support system can propose user role-specific measures to the users. Thus, users only receive suggestions for action and information within their scope.

Based on the different types of agents, a multi-agent framework is deducted. Figure 1 depicts a general set-up for a hierarchical MAS. The agents' action space represents measures, which the corresponding user or user group for each production resource can initiate within the scope of production control. In the example given, a second agent (e.g., a foreman) supervises two agents (e.g., machine operator). Between those two types of agents, the available action space and the observation space and, therefore, the available actions might differ. Therefore, an agent can be given a different responsibility depending on their respective abilities. For example, the agent representing a foreman might be able to change the volume of an order if necessary. In contrast, the machine operators might only determine the following order to be produced from a small number of orders.

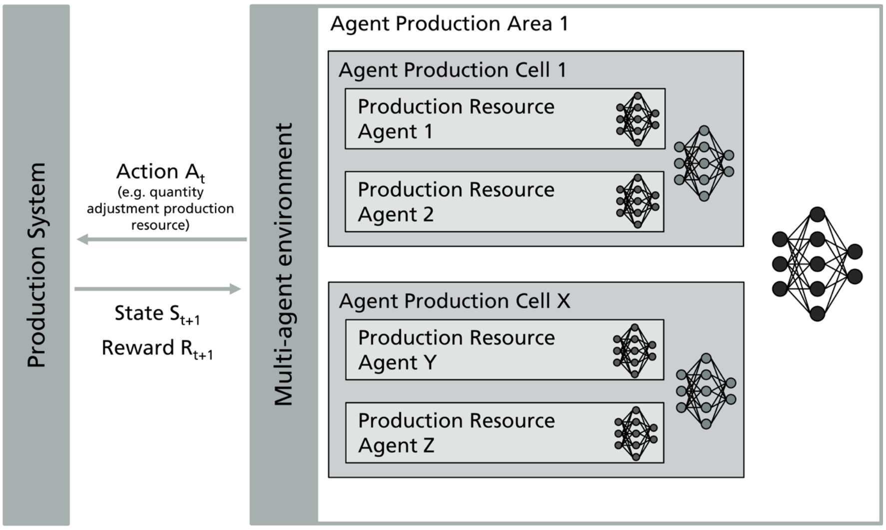


Figure 1: System architecture of the developed system with different types of agents

With the help of DRL, complex environments can be abstracted to select a suitable action given the situation [18]. Thus, enabling the determination of a policy for each separate agent, with regard to an individual observation space and action space. The agents' training occurs by interactions between the production system's simulation model and the MAS. Based on the reward given at the end of every episode (e.g., one week), the agents learn to choose the best possible action based on the system's current status. The trained agents can react near real-time on disruptions. Within the scope of this paper, PPO is used as a DRL algorithm because of its high robustness while exhibiting good learning behavior [31,32]. PPO uses the actor-critic approach, in which π is learned based on a value function [31]. This allows generalizing complex decision situations well [32]. Based on the multi-agent approach, it is possible to fulfill production control tasks while optimizing the logistic costs.

3.2 Explainability

The MAS with DRL presented in Section 3.1 focuses on improving decision-making concerning decision quality and reaction time. However, PPO is a black-box method using ANNs. In order to determine the influencing parameters on decisions, post-hoc methods are used in this paper. Thus, through DRL, a high decision quality can be maintained, and at the same time, the decision-making can be made explainable. To increase the users' confidence in the system, global explainability and local explainability are used to avoid unnecessary overrides. Post-hoc analysis requires the evaluation of the model as a whole. Therefore, it is necessary to evaluate every agent's ANN. For this purpose, all influencing factors are examined concerning their effects on selecting a particular action. Thus, an explainer can be generated, which shows the influencing factors for a given initial state and the resulting decision. Therefore, it is possible to equip a trained system with an explainer once, use them continuously and obtain a post-hoc explanation. In order to achieve this, the influence of the observations on the choice of action is calculated using SHAP values [27].

These SHAP values provide the results determined by the system. Starting from a base value (e.g., planned quantity), the influence of the individual input characteristics can be calculated to determine the resulting value [27]. SHAP enables both global and local explainability. With the help of global explainability, it is possible to present the decision-making process in a generally comprehensible way. Thus, relevant influencing parameters can be identified in general. On the one hand, this helps the user to understand the influences on the system's decision process. On the other hand, by determining the influence of different parameters, the agents' observation spaces can be adapted. This has a positive impact on the learning behavior of the agents and their ANN. The local explainability allows the determination of single influencing factors and their contribution to single decisions. Figure 2 schematically shows a so-called force plot. At the top, the specific action being explained can be seen. The features visualized in black led to a reduction from 70 to 40. The features in gray color show the features that stopped the reduction at 40. Thus, showing the influencing parameters' impact on the decision and to which extent a system built in a user-centric manner

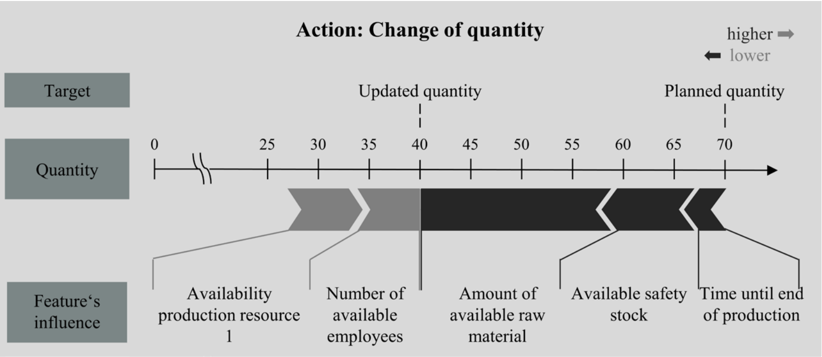


Figure 2: Decomposition of an individual action of one agents provides the influences for why this action was chosen in the given situation.

using MAS and leveraging advanced techniques such as DRL can use XAI to increase the users' trust in the proposed decisions while maintaining high decision quality and short reaction time.

4. Use Case

For the experiments being conducted within the scope of this paper, a simulation model of a production system has been used. The simulation model was built using Python as a programming language. Thus, realizing the short running times of the simulation model is beneficial for the number of iterations needed for the agents' learning process.

4.1 Simulation model

The simulation model is based on a real-life production system and consists of eight interlinked machines. Those eight machines represent a multi-stage production process. In total, the production process has four stages and parallel machines in two of them. Machines within the same stage have different (but sometimes overlapping) abilities regarding products, which can be processed as well as different processing times for specific products. Each product has to be processed once within each stage, and no skipping of process steps is allowed. Each machine can have four states (processing, waiting, set up, disturbed). Machine breakdowns are stochastically distributed and cannot be avoided. Orders for each episode (representing one week) are created based on a historic distribution at the start of each episode. Each order is characterized by product type and quantity. One episode contains, on average, 35 differing orders. All orders can be picked at the beginning of the episode. The set-up time needed for the production of an order is mainly based on the product type of the preceding order on the specific machine.

4.2 Multi-agent system

For the experiments conducted in the context of this paper, the focus has been on the improvement of order sequencing based on the system state. For this purpose, each production resource agent has up to five orders and the associated properties of the order (e.g., quantity, product type) in its observation space. Each unique order in the observation space is represented by an action in the agent's action space. Additionally to these actions, the agent can choose not to pick an order and wait if it is beneficial. If the previous action has been completed, agents can select a new action. The selection of actions from the agent's action space determines the sequence of production orders. During sequencing, further restrictions such as employee capacities, different processing times on different machines, and machine conditions (e.g., machine failures) have to be taken into account.

All agents receive a global reward at the end of the episode, determining how successful the past episode was. This approach has proven to be beneficial for the agents' learning behavior. Furthermore, local optima can be avoided while maximizing the reward [12]. Thus, resulting in better handling of the complexity of the simulation model. The reward consists of two parts; one part represents the logistical performance (e.g., lead time), and the other represents the logistical costs (e.g., inventory cost) that arise during production. If the lead time decreases, the logistical performance reward increases. If the inventory costs for an episode decrease, the logistical cost reward increases. A company-specific prioritization of the reward can be achieved by scaling the rewards.

The agents were trained by using PPO. The resulting ANNs were passed to SHAP Deep Explainer. Hereby, an explainer model could be built, allowing the decisions' explainability. This makes it possible to show both the individual decisions of the respective agents and the superordinate factors influencing the agent's decisions (global explainability). Particularly, the findings of global explainability were used to adjust the observation space and thereby improve the reward iteratively.

4.3 Results

In order to evaluate the performance of the developed MARL system, two conventional methods are used for comparison. In many companies, heuristics are common for sequencing orders within production control [33]. A widespread heuristic being used due to its simplicity is FIFO (First in - First out) [34]. The second heuristic used for the evaluation is “shortest set-up time next (SSTN)”.

For comparing FIFO, SSTN, and MARL, 52 episodes were simulated. All three approaches used the same initial production program. Figure 3 compares the average total reward and the corresponding components, which consist in this use case of lead time reward and inventory reward. For all episodes, the average total reward of MARL compared to FIFO was 22 % higher. Hereby, improvements were achieved to the same extent through increased logistical performance, represented by the lead time reward (by 24 %) as well as logistical cost, represented by the inventory reward (by 37 %). The average total reward of MARL compared to SSTN was 15 % higher. Thereby, the improvements stem from an improvement of the logistical cost (by 40 %), as well as improvements of the logistical performance (4 %). Local explainability is primarily intended for realizing user-centered information visualization within specific decision situations (e.g., selecting the following order). By identifying the importance of different features with the help of global explainability, the observation space can be adjusted. Thus, the MAS’s resulting reward can be increased.

MARL was able to increase the reward gained for production time as well as capital commitment costs compared to FIFO and SSTN, resulting in a higher overall reward. One challenge in optimizing logistic targets is to improve several metrics simultaneously. This arises because different objectives (e.g., short lead times, low inventory costs) interfere with each other. Optimizing those multiple objectives is complex. However, the developed approach, based on MARL, showed promising results of achieving a multi-objective optimization.

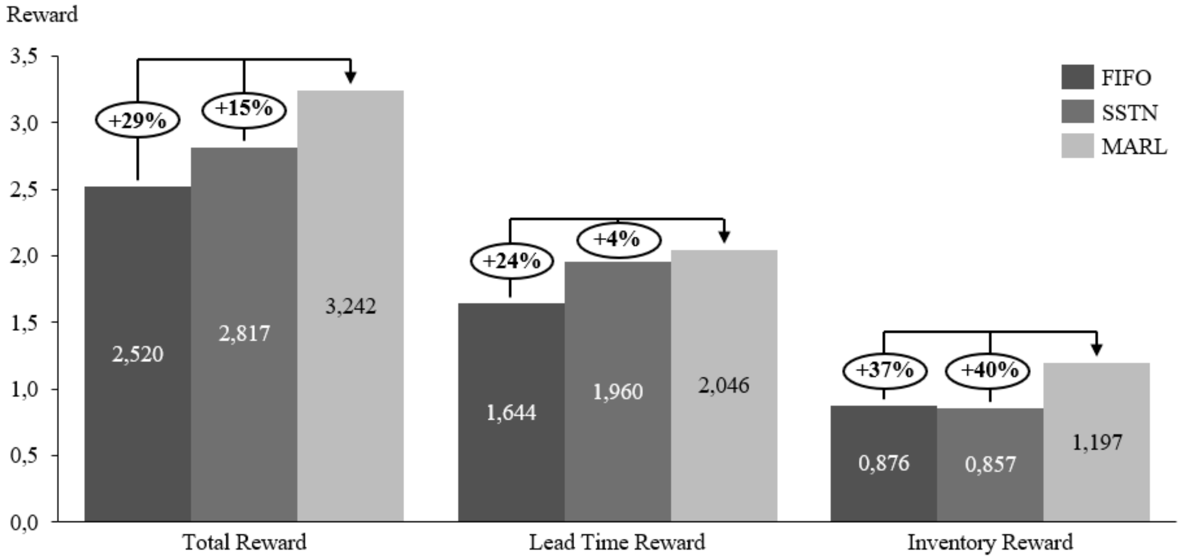


Figure 3: Comparison of the resulting rewards

5. Summary

With the help of the proposed approach, a production control based on MARL with explainable decisions can be realized. Deducted from different user roles, a MAS has been set up. With this work, it could be shown that a MAS with DRL offers the possibility to improve production control with regard to the defined reward. Combining a MAS, DRL, and XAI can improve decision quality and reaction time while also explaining the decisions being made. The focus on user-centricity is an essential component for the applicability. The current state of the XAI component indicates that it can enable increased trust in the AI system. Further investigations, especially concerning the use of local explainability, are necessary. Furthermore, the behavior of different DRL algorithms, besides PPO, will be tested. For further validation, the number of machines and products—and therefore the complexity of the production system— will be increased.

Acknowledgements

The project REIF – Resource-efficient, Economic, and Intelligent Food chain is funded by the German federal Ministry for Economic Affairs and Climate Action (BMWK).

References

[1] Reinhart, G., Zühlke, D., 2017. Von CIM zu Industrie 4.0, in: Reinhart, G. (Ed.), Handbuch Industrie 4.0. Carl Hanser Verlag GmbH & Co. KG, München, pp. XXXI–XXXIV.

[2] Schuh, G., Reinhart, G., Prote, J.-P., Sauermann, F., Horsthofer, J., Oppolzer, F., Knoll, D., 2019. Data Mining Definitions and Applications for the Management of Production Complexity. Procedia CIRP 81, 874–879.

[3] Cheng, Y., Chen, K., Sun, H., Zhang, Y., Tao, F., 2018. Data and knowledge mining with big data towards smart production. Journal of Industrial Information Integration 9, 1–13.

[4] Engelhardt, P.R., 2015. System für die RFID-gestützte situationsbasierte Produktionssteuerung in der auftragsbezogenen Fertigung und Montage. Diss. Techn. Univ. München, 2015. Utz, München.

[5] Blunck, H., Windt, K., 2013. Komplexität schafft Spielraum für Selbststeuerung. wt-online 2-2013 2, 109–113.

[6] Kletti, J., 2015. MES - Manufacturing Execution System. Springer Berlin Heidelberg, Berlin, Heidelberg.

[7] Lödding, H., Mundt, C., Winter, M., Heuer, T., Hübner, M., Seitz, M., Schmidhuber, M., Maibaum, J., Bank, L., Roth, S., Scherwitz, P., Theumer, P., 2020. PPS-Report 2019. TEWISS Verlag.

[8] Scherwitz, P., Bank, L., Roth, S., Theumer, P., Mundt, C., Winter, M., Heuer, T., Hübner, M., Seitz, M., Schmidhuber, M., Maibaum, J., 2020. Digitale Transformation in der Produktionsplanung und -steuerung. VT (4), 252–256.

[9] Usuga Cadavid, J.P., Lamouri, S., Grabot, B., Pellerin, R., Fortin, A., 2020. Machine learning applied in production planning and control: a state-of-the-art in the era of industry 4.0. Journal of Intelligent Manufacturing 31 (6), 1531–1558.

[10] Dittrich, M.-A., Fohlmeister, S., 2020. Cooperative multi-agent system for production control using reinforcement learning. CIRP Annals 69 (1), 389–392.

[11] Altenmüller, T., Stüker, T., Waschneck, B., Kuhnle, A., Lanza, G., 2020. Reinforcement learning for an intelligent and autonomous production control of complex job-shops under time constraints. Prod. Eng. Res. Devel. 14 (3), 319–328.

[12] Waschneck, B., 2020. Autonome Entscheidungsfindung in der Produktionssteuerung komplexer Werkstattfertigungen. Diss. Universität Stuttgart, 2020.

[13] Sutton, R.S., Barto, A.G., Barto, A., 2018. Reinforcement Learning: An introduction, 2nd ed. The MIT Press, Cambridge, MA, London, 526 pp.

[14] Burkart, N., Huber, M.F., 2021. A Survey on the Explainability of Supervised Machine Learning. jair 70, 245–317.

[15] Schuh, G., Stich, V. (Eds.), 2012. Produktionsplanung und -steuerung. 1: Grundlagen der PPS, 4th ed. Springer Vieweg, Berlin Heidelberg.

[16] Wiendahl, H.-P., 2014. Betriebsorganisation für Ingenieure, 8th ed. Hanser, München.

[17] François-Lavet, V., Henderson, P., Islam, R., Bellemare, M.G., Pineau, J., 2018. An Introduction to Deep Reinforcement Learning. FNT in Machine Learning 11 (3-4), 219–354.

[18] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529 (7587), 484–489.

[19] VDI/VDE Richtlinie 2653-1, 2018. Agentensysteme in der Automatisierungstechnik: Grundlagen. Beuth, Berlin, 24 pp.

[20] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M., 2013. Playing Atari with Deep Reinforcement Learning, 9 pp. http://arxiv.org/pdf/1312.5602v1.

[21] Roesch, M., Linder, C., Zimmermann, R., Rudolf, A., Hohmann, A., Reinhart, G., 2020. Smart Grid for Industry Using Multi-Agent Reinforcement Learning. Applied Sciences 10 (19), 6900.

[22] Murdoch, W.J., Singh, C., Kumbier, K., Abbasi-Asl, R., Yu, B., 2019. Definitions, methods, and applications in interpretable machine learning. Proceedings of the National Academy of Sciences of the United States of America 116 (44), 22071–22080.

[23] Holzinger, A., 2018. From Machine Learning to Explainable AI, in: 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), Kosice. IEEE, pp. 55–66.

[24] Gerlings, J., Shollo, A., Constantiou, I., 2021. Reviewing the Need for Explainable Artificial Intelligence (xAI), in: Proceedings of the 54th Hawaii International Conference on System Sciences.

[25] Ribeiro, M.T., Singh, S., Guestrin, C., 2016. Why Should I Trust You?, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA. ACM, New York, NY, USA, pp. 1135–1144.

[26] Shapley, L.S., 1953. 17. A value for n-Person Games, in: Arrow, K., Gale, D., Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games (AM-28), Volume II. Princeton University Press, Princeton, pp. 307–318.

[27] Lundberg, S., Lee, S.-I., 2017. A Unified Approach to Interpreting Model Predictions. http://arxiv.org/pdf/1705.07874v2.

[28] Rehse, J.-R., Mehdiyev, N., Fettke, P., 2019. Towards Explainable Process Predictions for Industry 4.0 in the DFKI-Smart-Lego-Factory. Künstl Intell 33, 181–187.

[29] Kuhnle, A., May, M.C., Schäfer, L., Lanza, G., 2021. Explainable reinforcement learning in production control of job shop manufacturing system. International Journal of Production Research 24 (4), 1–23.

[30] Huber, T., Weitz, K., André, E., Amir, O., 2021. Local and global explanations of agent behavior: Integrating strategy summaries with saliency maps. Artificial Intelligence 301, 103571.

[31] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017. Proximal Policy Optimization Algorithms, 12 pp. http://arxiv.org/pdf/1707.06347v2.

[32] Mayer, S., Classen, T., Endisch, C., 2021. Modular production control using deep reinforcement learning: proximal policy optimization. J Intell Manuf 32 (8), 2335–2351.

[33] Mönch, L. (Ed.), 2006. Agentenbasierte Produktionssteuerung komplexer Produktionssysteme, 1st ed. DUV Deutscher Universitäts-Verlag, s.l., 300 pp.

[34] Lödding, H., 2016. Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration, 3rd ed. Springer Vieweg, Berlin, Heidelberg, 664 pp..

Biography

Philipp Theumer (*1991) holds a Master's degree in Mechanical Engineering and Management from the Technical University of Munich. Since 2018 he is a research assistant at the Fraunhofer IGCV in the department of Production Management.

Florian Edenhofner (*1994) holds a Master's degree in Food Technology and Biotechnology from the Technical University of Munich. In 2021 he wrote his Master's Thesis at the Fraunhofer IGCV in the department of Production Management.

Roland Zimmermann (*1998) holds a B.Sc. in Computer Science from the University of Applied Sciences Augsburg. Since 2021 he has been enrolled at the University Augsburg for a Master's degree in Computer Science and is working at the Fraunhofer IGCV in the department of Production Management.

Alexander Zipfel (*1992) holds a Master's degree in Mechanical Engineering and Management from the Technical University of Munich. He leads the group for Production Planning and Control at the Fraunhofer IGCV in the department of Production Management.

3rd Conference on Production Systems and Logistics

Simulation Of All-solid-state Battery Manufacturing Routes*

Hans-Christoph Töpper[1,2], Maximilian Lechner[2] , Moritz Kupec[2], Rüdiger Daub[2]

[1]*TUMint. Energy Research GmbH, Lichtenbergstraße 4, 85748 Garching, Germany*

[2]*Institute of Machine Tools and Industrial Management, Technical University of Munich, Boltzmannstraße 15, 85748 Garching, Germany*

Abstract

All-solid-state batteries possess multiple advantages compared to already established battery technologies. Therefore, they are emerging as promising candidates for the sustainable storage of energy and acceleration of electrification in multiple fields. Despite the rapidly growing interest and vast material and cell design research activities, all-solid-state batteries are still in their infancy stage. Both market entry and full exploitation of the storage potential now mainly depend on the development of production technology empowering the large-scale breakthrough. Thereby, the selection of suitable manufacturing routes and development of production processes resulting from the novel components and materials plays a key role. To provide a better understanding and a systematic approach for the analysis of all-solid-state battery production, a holistic Matlab-based SimEvents factory simulation model is presented in this work. It enables the modeling and simulation of all-solid-state battery production scenarios consisting of a certain material choice, process steps, sequence, process parameters, storage capacity, and boundary conditions such as throughput, downtime, and scrap rate. An algorithm automatically performs the evaluation and comparison of the scenarios regarding production-related KPIs such as ramp-up time, capacity utilization, circulating stock, and storage load. In addition, the highly complex and nonlinear dependencies specific for all-solid-state battery production, as well as bottlenecks between the processes, are quantified. As a result, the factory model enables the optimization of manufacturing routes and production processes depending on the product design at a very early stage and the low-level maturity of this new energy storage technology.

Keywords

All-solid-state Battery Production; Factory Simulation; Production Optimization; Process Interdependencies

DOI: https://doi.org/10.15488/12152

---

* The paper successfully passed the peer review process but was not presented at a conference.

1. Introduction

In search of sustainable energy storage solutions suitable for countering the advancing climate change and providing energy to the steadily growing global population, various technologies are being pursued. Among these, the lithium-ion battery, in particular, has gained considerable importance in recent years resulting in a steep increase in global demand [1, 2]. According to forecasts annual sales figures might exceed 100 million by 2050 which is mainly driven by the rising popularity of electric vehicles [1, 2]. As a result, the annually produced storage capacity will reach the terawatt range in the near future [1, 2]. To meet this trend and since the lithium-ion battery is already technologically highly mature, solutions for increasing storage capacity must be found [3]. Currently, various approaches are being pursued aiming at further developing lithium-ion battery components or introducing fundamentally new battery technologies. With regard to the components, an increase in storage capacity is evoked, for example, by anode material compositions containing silicon [4, 5] and nickel-rich cathode structures [6, 7]. With regard to novel battery technologies, for example, the development of sodium-ion batteries [8, 9], lithium-air batteries [10, 11], and all-solid-state batteries [12, 13] is receiving strong boosts. Due to various advantages, the latter in particular seems to be a promising candidate for replacing conventional batteries. High energy and power densities [12, 14], improved safety [15, 16], and longer durability [14, 17] are just some of its potential key characteristics. However, from the point of view of the application in electric vehicles, the technology is still in its infancy stage. This is because all-solid-state battery cells are currently manufactured manually on a laboratory scale mainly for electrochemical characterization in formats not suitable for upscaling and with small-scale processes producing high scrap rates and costs [18, 19]. On a larger and industry-relevant scale, processes have not been established yet, but some thought has already been given to conceptualizing possible manufacturing processes [20, 21]. Accordingly, the current challenge is to realize the all-solid-state battery potential by developing production technology. To economically orientate the development of processes and manufacturing routes towards cost-effective production, an approach is presented in the following. Here, the technological focus is on sulfidic all-solid-state batteries.

2. Manufacturing strategies for sulfidic all-solid-state batteries

Sulfidic all-solid-state batteries consist of two electrodes and a separator and consequently do not differ from galvanic cells of conventional batteries [12, 22]. However, the materials applied in the components are new to battery technology. High capacity and low-density lithium metal is used as anode active material on a current collector for electrical contacting [12, 13]. The separator consists of densely packed ion-conducting solid sulfidic electrolyte particles held together by a polymeric binder [22, 23]. The composite cathode is composed of lithium-ion storing cathode active material particles, sulfidic solid electrolyte, and additives such as conductive carbon for electrical conductivity and binder for mechanical stability [23, 24]. For electrically contacting the composite cathode, these components are applied to a current collector. [23, 24] Figure 1 schematically depicts the structure of such a cell with its components. Since they do not play any further role in the following sections, the additives are not shown.

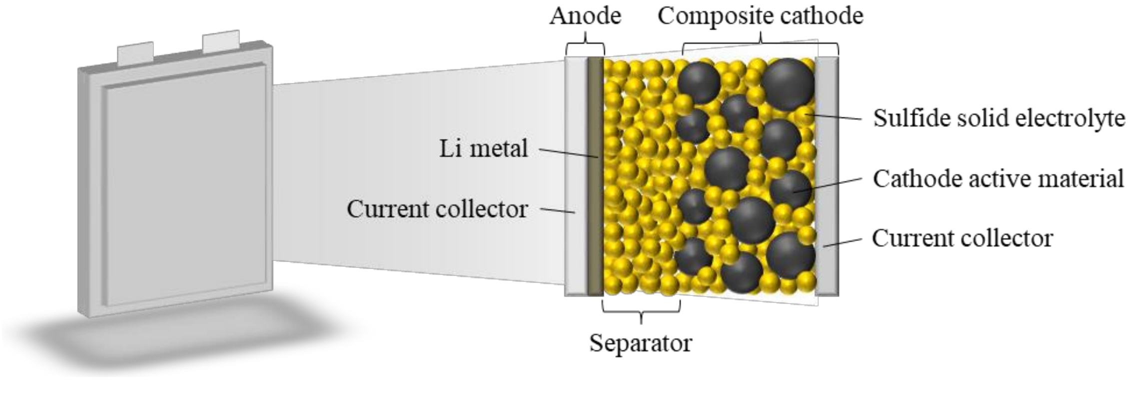


Figure 1: Composition of a sulfidic all-solid-state battery [14, 23]

According to current research, both composite cathode and separator can be manufactured in processes already used in the roll-to-roll production of conventional lithium-ion battery electrodes [19, 22, 25]. In the first step called mixing, the components of the separator and composite cathode are dispersed in a solvent. The result of this step is a slurry that is applied as a wet film to a substrate in the following coating step. Subsequently, the solvent is removed from the wet film in the drying step. The remaining porous dry film is then compacted in the final step using a calender. Since the performance of a sulfidic all-solid-state battery is determined by the contact of the solid-state particles, the compaction is of particular importance. Distinguishing the production from already established battery technologies is the fact that the process steps do not have to be carried out in parallel for composite cathode and separator. Theoretically, multiple combinations are conceivable with regard to the sequence. As an example, three of these strategies are shown in figure 2. According to current research, these are most promising [20, 23, 24].

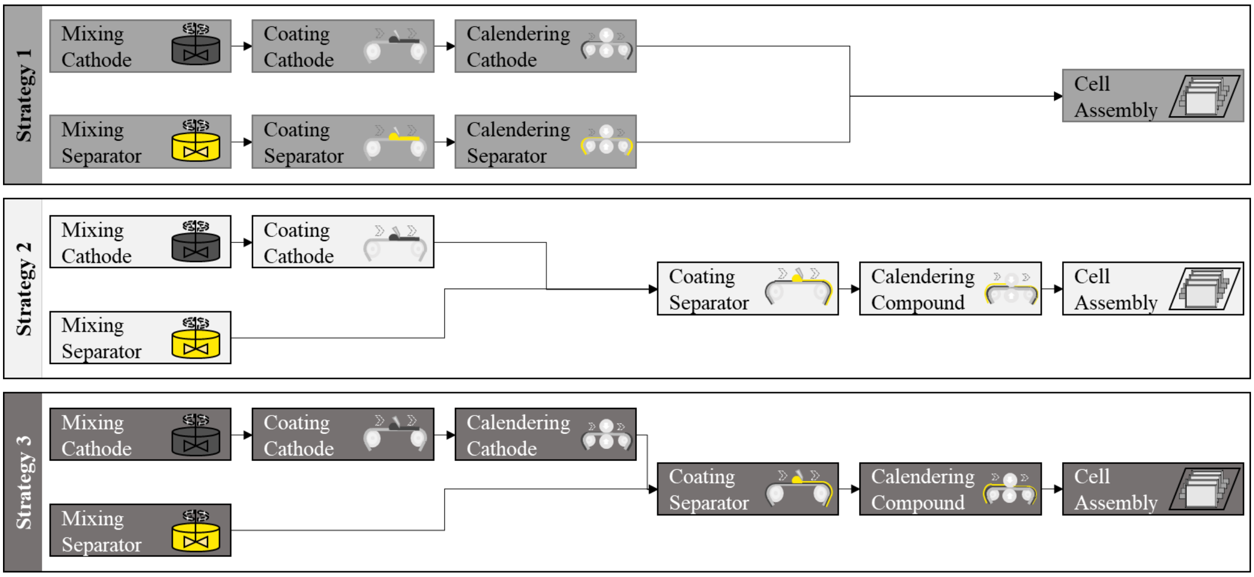


Figure 2: Examples of manufacturing strategies for the production of sulfidic all-solid-state batteries

The first manufacturing strategy is based on the production of conventional lithium-ion battery components, whereby the composite cathode and separator are manufactured in two separate process chains consisting of mixing, coating, drying, and calendering before being assembled to the cell. In the second strategy, the composite cathode slurry is mixed, coated onto the current collector, and then dried. Afterward, the separator slurry is applied to the cathode before the resulting compound is dried and compacted. By this type of processing, the separator penetrates the pores of the loosely packed composite cathode and only a single compaction step is performed for the compound. The third manufacturing strategy is similar to the second one, except that the separator is coated onto the already compressed composite cathode before the compound is dried and compressed. Consequently, no penetration of the separator slurry into the composite cathode occurs and two compaction steps are performed.

Based on these examples, it becomes clear that theoretically feasible manufacturing strategies for sulfidic all-solid-state batteries significantly differ regarding production technology. Therefore, the analysis of the process arrangements is made assessable by the presented approach aiming at identifying suitable strategies from an economic perspective.

3. Modeling of manufacturing routes

A simulation model quantifying the manufacturing strategy's impact on production-related key performance indicators (KPIs) and therefore enabling the economic evaluation was developed. The focus is on KPIs evaluating the efficiency of production such as throughput, ramp-up time, circulating stock, and storage utilization. The model enables the arrangement of processes according to the strategy and process parameters chosen to suit both material choice and battery cell format. To perform a simulation run and systematically analyze the influence of parameters, various boundary conditions can be chosen to define so-called production scenarios.

3.1 Simulation of production processes

The process steps mixing, coating, drying, and calendering posing the basic building blocks of each strategy are implemented in an event-based Simulink model in Matlab enabling a high degree of control over process parameters and design of interrelations between the individual steps. Both sequences as well as the number of processes and their parameters can be chosen by adding and connecting them on the Simulink interface. As an example, the model for the first manufacturing strategy in which composite cathode (top row) and separator (bottom row) are produced separately is shown in figure 3. Upon receiving a production order containing the material composition and required quantity, the individual process steps are executed. Their relationship is modeled via the ratios of the intermediate products. For example, the link between mixing and coating is described in terms of the volume of slurry needed to produce the required wet film possessing a certain length, width, and thickness. The process output of each step is by an intermediate storage uncovering bottlenecks and low capacity utilization. Once all process steps are completed, the simulated components are assembled into a cell for calculating the resulting storage capacity. This information is merged with all variables along the entire process chain at the end of each simulation run.

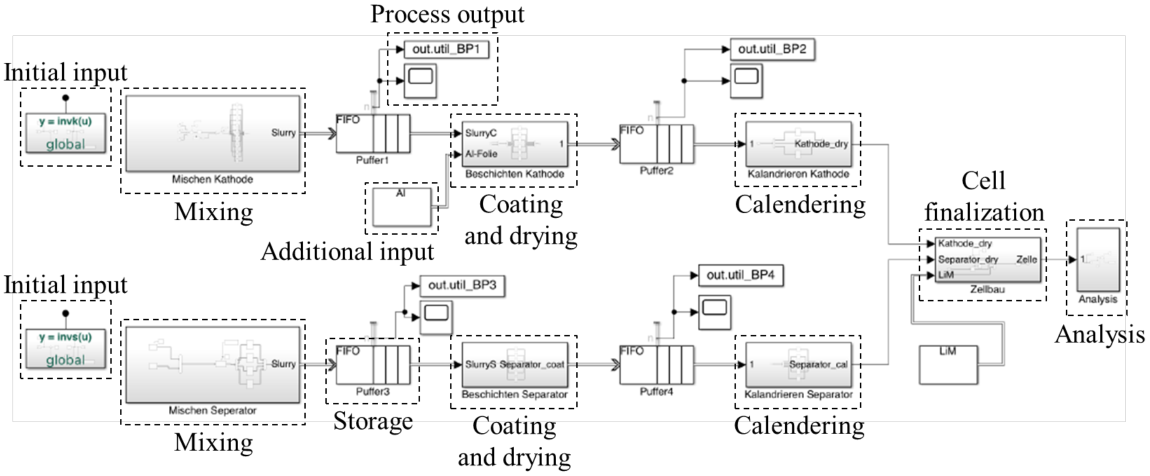


Figure 3: Model of the production processes in Matlab Simulink for manufacturing strategy 1

For each simulation run, various boundary conditions must be defined for the model. These include factors such as annual factory output, number of working days, and number of daily shifts as well as equipment effectiveness and scrap rate of each process step. Overall, this results in a tool supporting the systematic analysis of manufacturing strategies for producing sulfidic all-solid-state batteries.

4. Simulation of manufacturing strategies

To show the influence of the depicted manufacturing strategies on KPIs by applying the model, different production scenarios are defined. The shown selection can be extended by any strategies as well as boundary conditions or process parameters.

4.1 Definition of production scenarios

Different production scenarios for manufacturing the composite cathode separator compound are defined for each strategy varying equipment availability and scrap rate in three levels. For each simulation run, the annual storage capacity factory output is 20 GWh, work is carried out 300 days a year in three shifts each lasting eight hours. Likewise, the parameters of all process steps are kept constant for the isolated examination of availability and scrap rate. Table 1 contains the overview of the manufacturing strategies and parameter levels for each scenario resulting in a total of 27 simulation runs.

Table 1: Production scenarios for three different manufacturing strategies and variation of two parameters

Manufacturing strategy	Equipment availability in %	Scrap rate in %
1 (parallel)	100	0
2 (separator on porous cathode)	95	5
3 (separator on dense cathode)	90	10

The simulations are performed for a 150 mm by 150 mm pouch cell containing a total of 40 layers. The materials used are lithium metal as anode active material, NMC 622 as cathode active material, a sulfidic solid electrolyte, carbon fibers as conductive additive, and a polymeric binder. The electrode current collectors are made from aluminum and copper. Based on these assumptions, the calculations result in 739 Wh/l for the volumetric energy density of a cell is which is significantly higher compared to conventional batteries [26–28].

4.2 Analysis of throughput

The results from the simulation runs shown in figure 4 display that, given a certain availability and scarp rate along the process chain, the manufacturing strategy affects the number of cells produced per minute by up to 2 %. Differences in terms of throughput for strategy 1 compared to strategies 2 and 3 occur due to synchronization difficulties of the parallel process chains. The deviation increases as the equipment availability drops and the scrap rate declines. Thereby, the availability shows a much stronger impact on the throughput than the latter. This is due to the significant increase in idle times of intermediate products between processes and leads to a productivity drop of approximately 8 %. Consequently, from this point of view, it makes no difference whether the components are manufactured separately or the separator is coated directly onto the uncompressed or densified composite cathode.

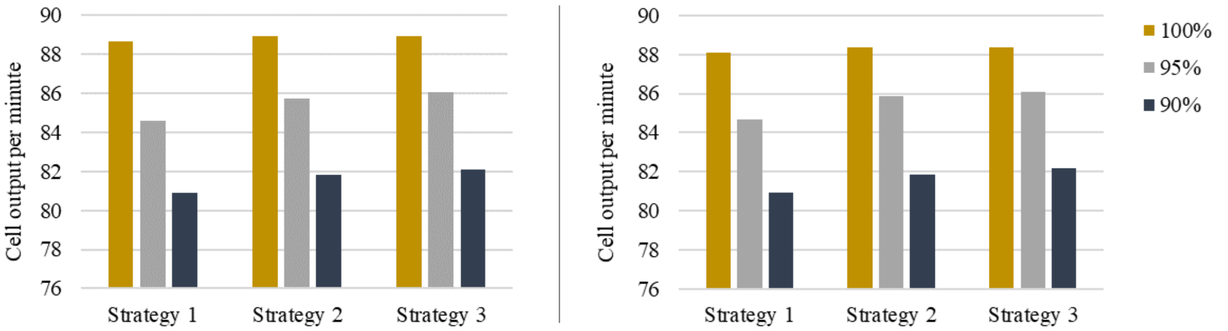


Figure 4: Influence of equipment availability on cell output per minute for 5 % (left) and 10 % (right) scrap rate

4.3 Analysis of circulating stock

A distinct pattern is emerging as the manufacturing strategy strongly influences the circulating stock along the process chain. Figure 5 shows a contrasting development of the number of composite cathode and separator batches. The circulating inventory of the composite cathode is lowest for strategy 1 and increases significantly for strategies 2 and 3. The differences of approximately 25 % and 50 % compared to strategy 1 are due to the increasing lengths of the process chains and the fact that manufacturing always starts with mixing and coating of the composite cathode. With regard to the separator, the trend is the opposite. The number of batches decreases when changing from strategy 1 to 2 or 3. This is due to the fact, that the separator is added to the compound at a later stage. However, the difference is not as distinct as for the composite cathode. Besides the manufacturing strategy, the equipment availability also affects the circulating stock.

This is because intermediate products such as slurries or dry coatings cannot be further processed as a result of higher process failure rates leading to fuller storages and therefore increased idle times. The scrap rate does not have a significant impact on this trend.

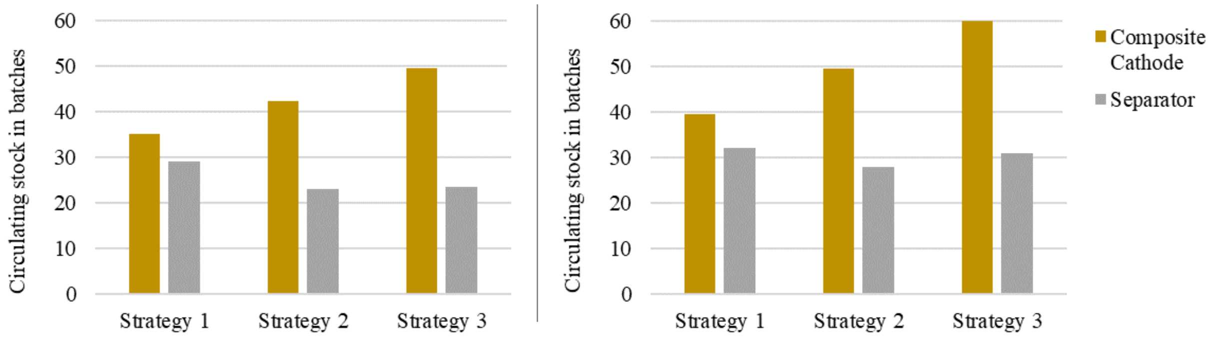


Figure 5: Circulating stock at 5 % scrap rate for 95 % (left) and 90 % (right) equipment availability

4.4 General comparison of manufacturing strategies

Besides providing a detailed analysis and explanation of the interaction between manufacturing strategies and production-related parameters, the approach mainly aims at the general comparison of processing routes. Spider diagrams show the comprised results of the process simulations, containing the normalized assessment on all KPIs as a function of the process availability. The higher the achieved rating for a certain KPI, the more advantageous the manufacturing strategy performs in this respect. Based on figure 6 showing the result for varying equipment availability at a scrap rate of 0 %, significant differences between the manufacturing strategies become apparent.

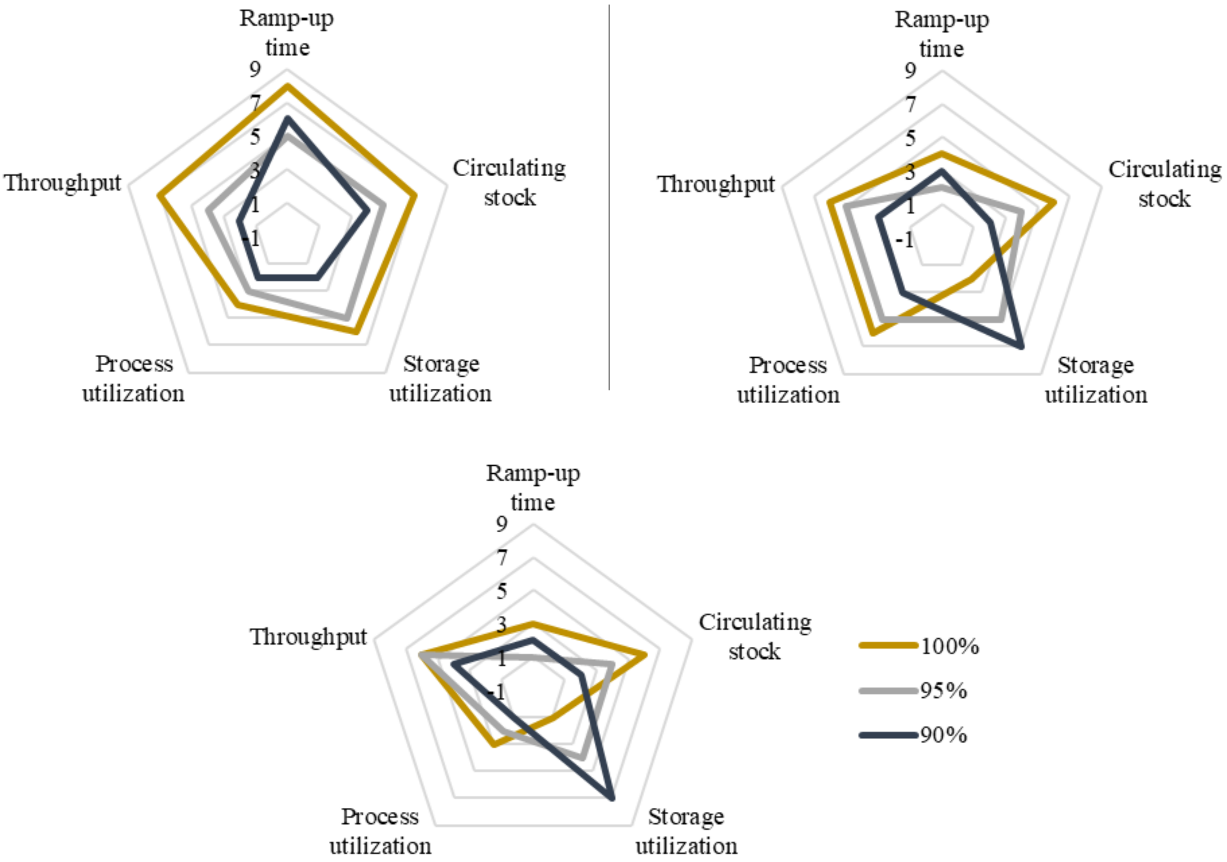


Figure 6: Holistic comparison of manufacturing strategies 1 (top left), 2 (top right), and 3 (bottom); increased relative advantage with increasing rating from 1 to 9

If, for example, the aim is the production of different cell formats with varying geometries or storage capacity, a strategy should be selected that has a short ramp-up time to achieve full utilization of the processes as quickly as possible. This requirement is best fulfilled by strategy 1. However, if the aim is to keep throughput as high as possible with fluctuating process availability, strategy 3 should be selected, as it shows the best performance in this respect. If the material costs tied up in processes and storage are to be kept as low as possible, strategy 2 or 3 should be chosen, as these have the lowest circulating stock. The evaluation of the key figures with regard to further settings can be carried out analogously.

5. Discussion

It turns out that the selection of a manufacturing strategy for sulfidic all-solid-state batteries from an economic point of view is not straightforward. The process model uncovers major differences in the performance of selected manufacturing strategies with respect to production-related KPIs. However, identifying the most advantageous strategy based on the presented results is not possible despite deep insights. This is substantiated by the fact, that each KPI will need to be weighed by resulting costs and their relation to the storage capacity. In addition, the components' performance regarding aspects such as rate capability and long-term durability need to be considered as well. Also, the overarching strategy of the cell producer which might either lean toward flexibility or throughput maximization will have to be taken into account. Thus, the model is just one of many building blocks in the decision-making process aiming at establishing a cost-effective and high-quality production of all-solid-state battery cells.

6. Conclusion and outlook

For the production of sulfidic all-solid-state batteries and utilization of their potential, a suitable manufacturing strategy must be chosen. As various processing routes are conceivable at the present stage of the technology, there is a need for an economic evaluation of these streamlining the cost-effective process development. To perform this, a Matlab-based model was developed providing the relative comparison of manufacturing strategies regarding production-related KPIs. Its operation was demonstrated by comparing three manufacturing strategies. In the first case, the composite cathode and separator are produced in two parallel process chains consisting of mixing, coating, drying, and calendering. In the second case, the separator is coated onto the uncompressed composite cathode, and in the third case, the separator slurry is applied to a densified cathode. The simulation-based analysis of the three strategies as a function of scrap rate and equipment availability shows differing routes, in some cases significantly, with regard to production-related KPIs such as throughput and circulating stock. As a result, this contributes to estimating the effects of the chosen processing procedure on the later to be established industrial production.

Future works and examinations will further analyze various manufacturing strategies. For this purpose, uncertainties regarding the processes and their parameters will be integrated into the model. In addition, the production steps of cell assembly and finalization will be modeled. Once this is completed, the authors aim at monetarily weighting the evaluation criteria to support the selection more clearly in this respect.

Acknowledgement

The authors express their sincere gratitude to the Bavarian Ministry of Economic Affairs, Regional Development and Energy for the funding of the research project "Industrialisierbarkeit von Festkörperelektrolytzellen".

References

[1] Xu, C., Dai, Q., Gaines, L., Hu, M. *et al.*, 2020. Future material demand for automotive lithium-based batteries *1*, p. 437.

[2] Zhang, J., Wang, Z., Liu, P., Zhang, Z., 2020. Energy consumption analysis and prediction of electric vehicles based on real-world driving data *275*, p. 115408.

[3] Choi, S., Wang, G., 2018. Advanced Lithium-Ion Batteries for Practical Applications: Technology, Development, and Future Perspectives *3*, p. 1700376.

[4] Chen, X., Li, H., Yan, Z., Cheng, F. *et al.*, 2019. Structure design and mechanism analysis of silicon anode for lithium-ion batteries *62*, p. 1515.

[5] Franco Gonzalez, A., Yang, N.-H., Liu, R.-S., 2017. Silicon Anode Design for Lithium-Ion Batteries: Progress and Perspectives *121*, p. 27775.

[6] Kim, J., Lee, H., Cha, H., Yoon, M. *et al.*, 2018. Prospect and Reality of Ni-Rich Cathode for Commercialization *8*, p. 1702028.

[7] Myung, S.-T., Maglia, F., Park, K.-J., Yoon, C.S. *et al.*, 2017. Nickel-Rich Layered Cathode Materials for Automotive Lithium-Ion Batteries: Achievements and Perspectives *2*, p. 196.

[8] Hwang, J.-Y., Myung, S.-T., Sun, Y.-K., 2017. Sodium-ion batteries: present and future. Chem Soc Rev *46*, p. 3529.

[9] Liu, Q., Hu, Z., Li, W., Zou, C. *et al.*, 2021. Sodium transition metal oxides: the preferred cathode choice for future sodium-ion batteries? *14*, p. 158.

[10] Geng, D., Ding, N., Hor, T.S.A., Chien, S.W. *et al.*, 2016. From Lithium-Oxygen to Lithium-Air Batteries: Challenges and Opportunities *6*, p. 1502164.

[11] Girishkumar, G., McCloskey, B., Luntz, A.C., Swanson, S. *et al.*, 2010. Lithium−Air Battery: Promise and Challenges *1*, p. 2193.

[12] Janek, J., Zeier, W.G., 2016. A solid future for battery development *1*, p. 1167.

[13] Varzi, A., Raccichini, R., Passerini, S., Scrosati, B., 2016. Challenges and prospects of the role of solid electrolytes in the revitalization of lithium metal batteries *4*, p. 17251.

[14] Kato, Y., Hori, S., Saito, T., Suzuki, K. *et al.*, 2016. High-power all-solid-state batteries using sulfide superionic conductors *1*, p. 652.

[15] Luntz, A.C., Voss, J., Reuter, K., 2015. Interfacial challenges in solid-state Li ion batteries *6*, p. 4599.

[16] Wu, F., Fitzhugh, W., Ye, L., Ning, J. *et al.*, 2018. Advanced sulfide solid electrolyte by core-shell structural design. Nat Commun *9*, p. 4037.

[17] Froboese, L., Groffmann, L., Monsees, F., Helmers, L. *et al.*, 2020. Enhancing the Lithium Ion Conductivity of an All Solid-State Electrolyte via Dry and Solvent-Free Scalable Series Production Processes *167*, p. 20558.

[18] Ito, S., Fujiki, S., Yamada, T., Aihara, Y. *et al.*, 2014. A rocking chair type all-solid-state lithium ion battery adopting Li2O–ZrO2 coated LiNi0.8Co0.15Al0.05O2 and a sulfide based electrolyte *248*, p. 943.

[19] Kim, M.-J., Park, J.-W., Kim, B.G., Lee, Y.-J. *et al.*, 2020. Facile fabrication of solution-processed solid-electrolytes for high-energy-density all-solid-state-batteries by enhanced interfacial contact. Sci Rep *10*, p. 11923.

[20] Duffner, F., Kronemeyer, N., Tübke, J., Leker, J. *et al.*, 2021. Post-lithium-ion battery cell production and its compatibility with lithium-ion cell production infrastructure *6*, p. 123.

[21] Tan, D.H.S., Banerjee, A., Chen, Z., Meng, Y.S., 2020. From nanoscale interface characterization to sustainable energy storage using all-solid-state batteries. Nat Nanotechnol *15*, p. 170.

[22] Sakuda, A., Hayashi, A., Ohtomo, T., Hama, S. *et al.*, 2011. All-solid-state lithium secondary batteries using LiCoO2 particles with pulsed laser deposition coatings of Li2S–P2S5 solid electrolytes *196*, p. 6735.

[23] Ates, T., Keller, M., Kulisch, J., Adermann, T. *et al.*, 2019. Development of an all-solid-state lithium battery by slurry-coating procedures using a sulfidic electrolyte *17*, p. 204.

[24] Sakuda, A., Kuratani, K., Yamamoto, M., Takahashi, M. *et al.*, 2017. All-Solid-State Battery Electrode Sheets Prepared by a Slurry Coating Process *164*, A2474-A2478.

[25] Kwade, A., Haselrieder, W., Leithoff, R., Modlinger, A. *et al.*, 2018. Current status and challenges for automotive battery production technologies *3*, p. 290.

[26] Li, J., Du, Z., Ruther, R.E., AN, S.J. *et al.*, 2017. Toward Low-Cost, High-Energy Density, and High-Power Density Lithium-Ion Batteries *69*, p. 1484.

[27] Manthiram, A., 2017. An Outlook on Lithium Ion Battery Technology. ACS Cent Sci *3*, p. 1063.

[28] Placke, T., Kloepsch, R., Dühnen, S., Winter, M., 2017. Lithium ion, lithium metal, and alternative rechargeable battery technologies: the odyssey for high energy density *21*, p. 1939.

Biography

Hans-Christoph Töpper graduated in mechanical engineering at the Karlsruhe Institute of Technology and is now working as a research associate at the TUMint. Energy Research GmbH and Institute for Machine Tools and Industrial Management (*iwb)* at the Technical University of Munich.

Maximilian Lechner completed his studies in mechanical engineering at the Technical University of Munich in 2021 and is currently working as a research associate in the department of battery production at the Institute for Machine Tools and Industrial Management.

Moritz Kupec studies mechanical engineering (M.Sc.) with a focus on production and logistics at the Technical university of Munich. He has been working as a research assistant in the department of battery production at the Institute for Machine Tools and Industrial Management since 2019.

Rüdiger Daub has been head of the Chair of Production Technology and Energy Storage Systems at *iwb* at Technical University of Munich and at the Fraunhofer Institute for Foundry, Composites and Processing Technology IGCV since 2021. Previously Prof. Dr.-Ing. Rüdiger Daub has held several positions at BMW Group Munich, most recently as head of technology development and prototype construction for lithium-ion cells.

3rd Conference on Production Systems and Logistics

Augmented Virtuality Data Annotation And Human-In-The-Loop Refinement For RGBD Data In Industrial Bin-Picking Scenarios

Andreas Blank[1], Lukas Baier[1], Maximilian Zwingel[1], Jörg Franke[1]
[1]Institute for Factory Automation and Production Systems, FAU Erlangen-Nürnberg, Germany

Abstract

Beyond conventional automated tasks, autonomous robot capabilities aside to human cognitive skills are gaining importance. This comprises goods commissioning and material supply in intralogistics as well as material feeding and assembly operations in production. Deep learning-based computer vision is considered as enabler for autonomy. Currently, the effort to generate specific datasets is challenging. Adaptation of new components often also results in downtimes. The objective of this paper is to propose an augmented virtuality (AV) based RGBD data annotation and refinement method. The approach reduces required effort in initial dataset generation to enable prior system commissioning and enables dataset quality improvement up to operational readiness during ramp-up. In addition, remote fault intervention through a teleoperation interface is provided to increase operational system availability. Several components within a real-world experimental bin-picking setup serve for evaluation. The results are quantified by comparison to established annotation methods and through known evaluation metrics for pose estimation in bin-picking scenarios. The results enable to derive accurate and more time-efficient data annotation for different algorithms. The AV approach shows a noticeable reduction in required effort and timespan for annotation as well as dataset refinement.

Keywords

Data Annotation; Augmented Virtuality; Human-in-the-Loop; Machine Learning; Bin-Picking

1. Introduction

Short product life cycles, an increasing amount of product variants and more complex goods pose challenges to the manufacturing industry. Flexible automation, involving robot systems, contribute to improve the situation. However, conventional automation reaches limitations in scenarios with uncertainties, including industrial bin-picking for material supply, machine feeding and assembly. Deep learning (DL) is an enabler for autonomous robot capabilities able to cope with such complex tasks [1]. Yet, the required dataset generation is time-consuming and thus costly [2]. In addition, downtimes may occur on system ramp-up [3].

In this context, the contribution of this paper is an augmented virtuality (AV)-based real-world RGBD data annotation and human-in-the-loop (HuITL) dataset refinement method. Objectives of the method are to reduce the effort and time spent in initial dataset generation for industrial bin picking and tuning the dataset up to operational readiness during ramp-up phase. In a single operation, both data for object classification and localization as well as data for 6DoF pose estimation are annotated. When in online refinement mode, in addition, the cognitive skills of the human remote operator are exploited to provide fault intervention and proper task solution for autonomous handling from distance by a teleoperation interface. The method enables accurate and more time-efficient RGBD data annotation and refinement. Thereby a noticeable reduction in required effort for successful application deployment and adaptation in industrial bin picking is achieved.

DOI: https://doi.org/10.15488/12184

2. Related Work: Data Generation and Human-In-The-Loop in Industrial Bin-Picking

In this section, progresses made in data generation for DL-based object recognition and regarding HuITL approaches for improving autonomous robot skills within industrial bin-picking applications are reviewed.

2.1 Data Generation for Object Recognition

The classification of algorithms can be done by different criteria, such as the image processing or input data. This includes object classification, localization, segmentation and pose estimation as well as solving combinations of these. As input, 2D-RGB, depth data or both, as well derived point clouds are common [4].

Although DL proves to outperform traditional algorithms, efforts required for specific dataset generation and training parametrization are still high [2]. A large quantity of data is required to improve the performance as well as to reduce the risk of overfitting. Data itself proves to be both the constraining and the driving factor. For data generation multiple techniques exist (cf. Figure 1). Depending on the type of input data, annotation is performed by 2D- and 3D-bounding boxes, 6DoF pose specification or pixel-wise labeling.

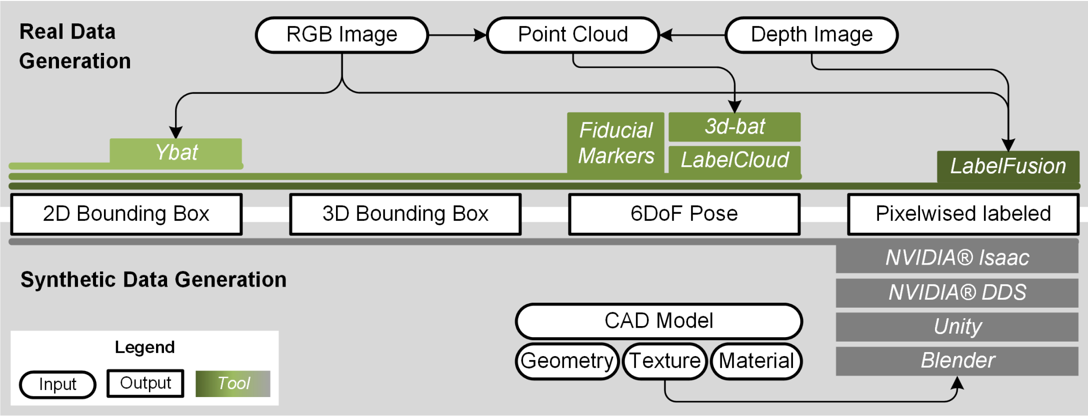


Figure 1: Data annotation types (center), procedures for annotating real-world data with common tools (top) as well as procedures and tools for automated generation of synthetic data (bottom)

Visual fiducial markers are state of art to determine ground truth pose for real-world objects [5]. Utilizing markers, however, is cumbersome and time-consuming. Components difficult to annotate exist due to geometric complexity. In addition, occlusions in multi-object scenarios complicate marker application.

For annotating objects based on recorded sensor data, software such as YOLO BBox Annotation Tool (Ybat), 3d-bat [6] and labelCloud [7] are important tools. Further available point cloud labeling tools focus mainly on autonomous driving [8]. Another solution is LabelFusion [9], a pipeline utilizing recorded RGBD video-data to produce pixel-wise object masks and 6DoF object pose labels. For software tools it is common to perform labeling utilizing computer monitor based graphical user interfaces (GUI). Thereby, components are iteratively aligned through definition of distinctive object related points and subsequent object boundaries are manually defined. For LabelFusion, in addition, a preceding 3D reconstruction step is required. Current solutions for labeling are time-consuming and common GUIs lack in spatial scene representation, especially complicating 6DoF pose annotation. In contrast, VR shows potential for efficient labeling. While VR solutions exist for semantic segmentation of large landscapes [10], approaches for appropriate labeling of data required for accurate robotic grasping operations are not available.

Synthetic data represent an alternative requiring less effort [11]. Thereby, an increased variance of 6DoF poses and camera perspectives as well as control of the image rendering are achievable. Digital component models involving texture and material specifications as well as current R&D-results on generative adversarial networks, domain adaption and domain randomization enable a more realistic dataset generation. However, closing the domain gap between synthetic and real-world data is still challenging [12].

2.2 Human-in-the-Loop Intervention and Dataset Refinement

Although datasets are generated component-specific, fault incidents of autonomous robots still occur. Since a machine operator may not always be nearby or on-site fault clearance may be harmful, effective remote intervention solutions are necessary. Current research focusses on teleoperated intervention [3]. However, relying solely on remote manipulation does not enable system adaptation. This is a relevant aspect, since root causes of failures often recur as well as operator time is expensive. Once a model is trained, human cognitive skills remain valuable to interpret and review model predictions as well as to enable dataset refinement [13,14]. Yu et al. show the potential of iterative refinement, especially in a more application-specific manner [15]. However, in industrial bin-picking exist a lack of HuITL dataset refinement methods.

3. Augmented Virtuality Data Annotation and Human-in-the-Loop Refinement and Intervention

Following the AV based real-world data annotation and HuITL refinement and intervention are described. First, an overview on architectural level is given (cf. Figure 2), followed by a description of the method.

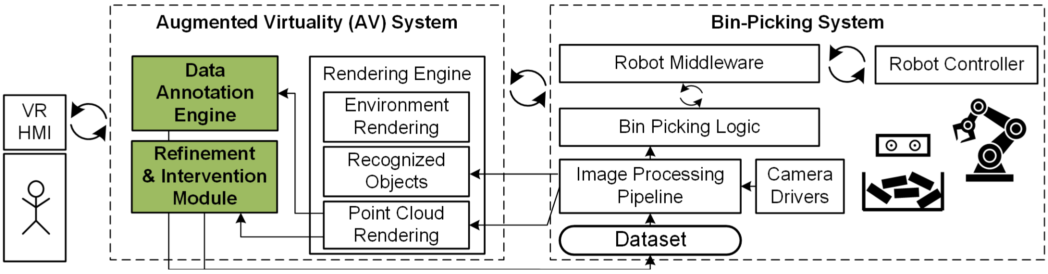


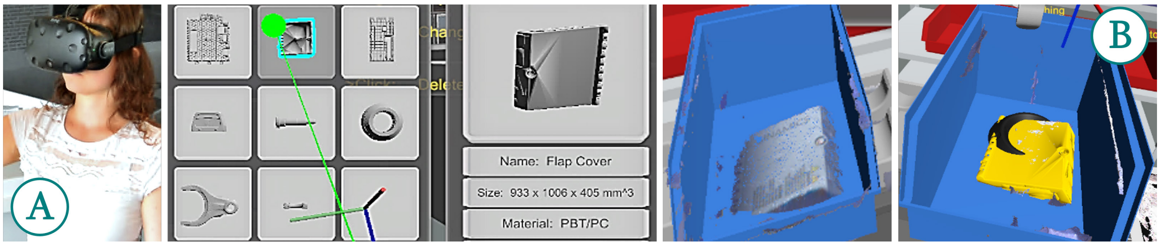
Figure 2: System overview as schematic technical architecture modelling (TAM)

3.1 Overview: Schematic Technical Architecture Modelling (TAM)

The AV rendering engine (cf. Figure 2) is described in previous work [16,17]. It involves rendering of the known environment, recognized objects as well as rendering of a pre-segmented point cloud. The illustrated image processing pipeline serves both systems for component classification, localization and 6DoF pose estimation. The exemplary utilized pipeline within this work is described in [4]. A robot control middleware is required to automate parts of the data annotation process (e. g. by changing camera perspective) as well as for HuITL intervention. The utilized middleware based on the Robot Operating System (ROS) is available open-source and is described in previous work [18]. Regarding the bin-picking system, the method requires access to the DL-dataset. At least weights should be interchangeable. To achieve a modular bin-picking system, a skill based logic using ROS actions (e. g. for motion control) is implemented [19]. Thereby, triggering and composition of skills as orchestration to a bin picking task is performed by a state machine.

3.2 Data Annotation Engine

The data annotation (Figure 2) is structured in: system parameterization and model selection, labeling and data generation (3.2.1) as well as automated data complementation (3.2.2). Although the descriptions address online data processing, the method is suitable for data generation based on once recorded offline sensor data.



Name: Flap Cover
Size: 933 x 1006 x 405 mm^3
Material: PBT/PC

Figure 3: AV Annotation: model selection (A) and labeling process (B) for data annotation

3.2.1 System Parameterization and Model Selection as well as Labeling and Data Generation

According to the bin-picking system, specifically the type of camera input data, the corresponding resolution and the type of desired output data (e. g. color images, point cloud, segmentation map, etc.) are determined. The camera transformation $\mathbf{T}_{\text{world}}^{\text{camera}}$ with respect to the world frame is calculated by the AV.

Within the annotation engine, new or not recognized components are rendered using their point cloud. The objective is to label these with matching models. For this purpose, the corresponding model is initialized (cf. Figure 3 (A)). The transformation $\mathbf{T}_{\text{world}}^{\text{object}}$ of components with respect to the AV world frame is determined within initialization step as well as the transformation $\mathbf{T}_{\text{camera}}^{\text{object}}$ is calculated. Thus, through visual alignment of the virtual component within the point cloud, the required ground truth is obtained (cf. Figure 3 (B)).

Since the coordinate system in rendering engines are based left-handed, whereas the camera coordinate system in robotics is often right-handed, transformations are applied. Finally, the labeling results are stored.

3.2.2 Automated Robot-Assisted Data Complementation

To further reduce effort in data generation, an optional automated robot-assisted data complementation is provided. This is applicable for setups with a robot-wrist mounted camera. Due to the previous labeling, the 6DoF poses of objects in space are known. Subsequently, the robot arm is remotely manipulated by the operator or automatically moved according to a pre-defined trajectory above and along the components for gaining further perspective object views. Thereby, new real-world data is annotated automatically.

3.3 Refinement and Intervention Module (HuITL)

Whereas AV annotation is employed for initial data generation, the refinement module (Figure 2) addresses system ramp-up enabling continuous dataset improvement. The method is structured in: interconnection and environment rendering (see 3.3.1), fault clearance (3.3.2) as well as annotation and refinement (3.3.3).

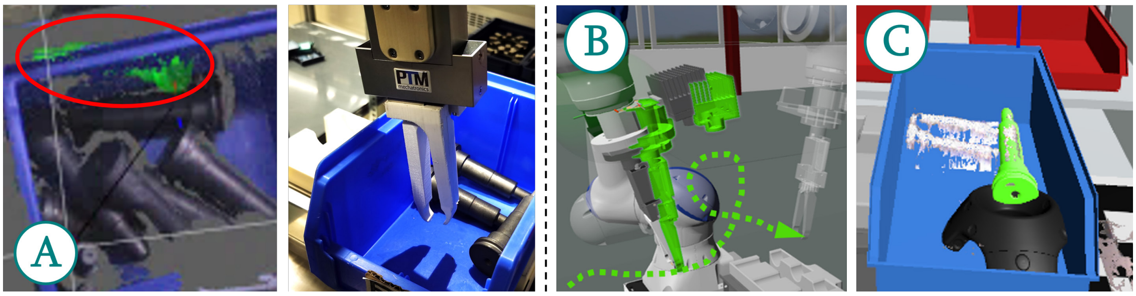


Figure 4: HuITL refinement and intervention: fault scenario caused by low confidence pose estimate (A), schematic illustration of teleoperated motion control (B) as well as hybrid component pose determination and annotation (C)

3.3.1 Interconnection and Environment Rendering

Once a fault is indicated by a linked system, an operator interconnects via the AV system. Subsequent, the robot environment involving system models and camera data is rendered. The fault condition itself might be triggered through a state-machine of the bin-picking system, caused by a repeatedly incorrect or low-confidence pose estimation as well as by multiple failed grasping attempts (cf. Figure 4 (A)).

3.3.2 Fault Clearance

The intervention module initially visualizes the last pose estimate (cf. Figure 4 (A)). In case a robot movement is required for solving (e. g. for component grasping or to realign camera, a component or the manipulator), the teleoperation interface is utilized (B). If the failure is caused by component localization or pose estimation, operators determine a reasonable solution (e. g. by target pose specification) based on the corresponding object model (C). This serves the bin-picking system as input for automated task completion.

3.3.3 Data Annotation and Network Refinement

The provided solution by the operator is stored as input for dataset refinement. The procedure is identical to AV annotation process. It differs in not generating an initial dataset, instead extending the existing dataset with additional more application-specific labels. The network model is updated based on the pre-trained model. Thereby, the latent network parameters are optimized towards the bin-picking application scenario.

4. Setup and Procedure of Experiments

The system setup, design of experiments as well as evaluation metrics are described in the following sections.

4.1 Demonstrator System Setup

To evaluate the method, a bin-picking testbed is implemented (cf. Figure 5 (A)). Visual perception is performed by a roboception rc_visard 65 stereo camera. The camera is mounted at the robot wrist. The camera delivers a depth image with a resolution of 640 × 480 pixels at a frame rate of 25 Hz. The highest available depth image resolution of 1280 × 960 pixels is not utilized due to lower frame rates. The camera depth deviation is specified with ±0.5 mm at 200 mm object distance and ±15 mm deviation at 1000 mm. Within the testbed a Yaskawa HC10 articulated robot is used. The system is based on ROS Melodic and Unity3D. ArUco markers are utilized for calibration between robot, camera and workspace.

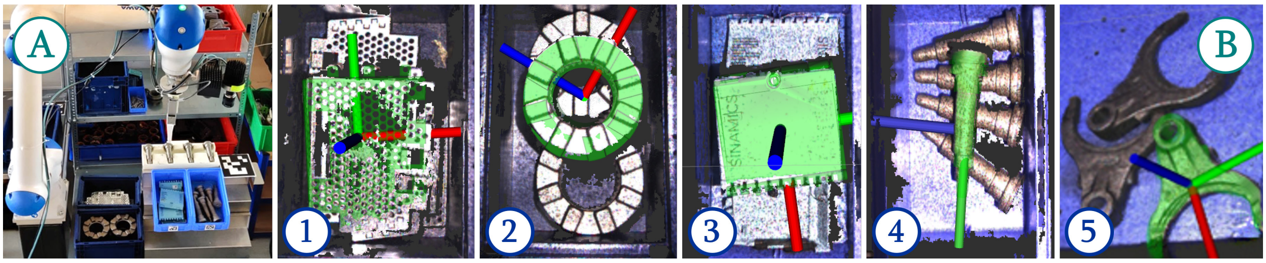


Figure 5: Bin-picking testbed (A) with evaluation components (B) – I/O shield (1), shifting sleeve (2), flap cover (3), shifting rod (4) and shifting fork (5); qualitative visual pose estimation results based on evaluation components (B).

The components under investigation are metallic and non-metallic parts of differing characteristics (i. a. material, shape, texture and surface) and will be referred to as: I/O shield (1), shifting sleeve (2), flap cover (3), shifting rod (4) and shifting fork (5).

Regarding materials, the flap cover (3) is made of polymer blend while the other parts (1,2,4,5) are metal. Geometries are a mixture of cylindrical and planar parts. Additionally, flap cover is matt and slightly textured while all metal parts are texture-less. I/O shield (1), in addition, challenges with a perforated surface.

4.2 Procedure of Experiments

Three experiments are performed for evaluation. First, the required timespan and the accuracy achieved when labeling are assessed. Here AV annotation is compared to the established AprilTag marker technique, utilized as standard for real-world data annotation in 6DoF pose estimation as well as been proven for high accuracy regarding annotation output [4,5]. This experiment serves to compare the performance of the proposed method in terms of accuracy and time effort to the state of the art. Further, to demonstrate feasibility and successful operability of the annotation outcomes provided by the proposed method, training results are tested by an image processing pipeline involving YOLOv3 [20] for object localization as well as Frustum PointNets [21] (FPN) for pose estimation within the described real-world bin picking testbed (Figure 5). Here, achieved object classification and localization success rates as well as 6DoF pose estimation accuracies obtained are assessed based on established evaluation metrics (cf. 4.3). Finally, a small user study is carried out to determine the Mean Time to Repair (MTTR) of the proposed teleoperation based fault intervention.

Figure 6: Ground truth determination (A) – (C) as well as technique used for preparing AprilTag comparison (D) – (F)

For accuracy evaluation, ground truth is required. First, ArUco markers serve for calibration (Figure 6 (A)). Afterwards, the test component is placed upon the ArUco marker and aligned (B). At the same time, the component is configured within a simulation (C). As a result digital and real-world components coincide. Finally, the ground truth 6DoF pose acquired is stored. Several steps are required for data annotation via AprilTag markers ((D)-(F)). First, for preparation the component must be measured to obtain center of gravity and surface middle (D). Subsequent, two markers are placed, one as reference within the scene and another in the component surface middle (E). Specifications obtained are stored for parametrization. Finally, the pose, the RGB image and the point cloud are recorded (F).

For comparison, AV-based annotation and AprilTag both are performed under consideration of the main influencing factors: namely the person conducting the annotation and the component to be annotated. For this purpose, a user study is carried out, involving five volunteers with annotation experience. Thereby, the five components ((1)-(5)) are labelled. As resulting measures, the timespan required for preparing and conducting labeling is recorded. The achieved accuracy is evaluated according to the metrics described (4.3).

To determine the MTTR, the study is extended. First, two failure types are defined – low-confidence pose estimation and failed grasping. A random selection between both is performed for each intervention. Within a session, the timespan between the following steps is measured for evaluation: putting on the VR headset, interconnection, spatial acquisition of the fault case, teleoperating the robot and target pose determination.

4.3 Evaluation Metrics

Throughout the literature, several different metrics for evaluating 6DoF pose accuracy have been proposed. To quantify the labeling as well as the resulting pose estimation accuracy, the following metrics are chosen.

Average Distance (ADD) [22] computes average distance between ground truth 6DoF pose and labeled or estimated pose utilizing the component model $\mathcal{M}$. With given ground truth pose $\widehat{\mathbf{P}}$ and estimated pose $\overline{\mathbf{P}}$, the average distance of model points is calculated (cf. [23]) as – utilized for test components (1), (3) and (5).

$$e_{ADD}(\widehat{\boldsymbol{P}},\overline{\boldsymbol{P}};\mathcal{M}) = \underset{x\in\mathcal{M}}{avg} \left\| \overline{\boldsymbol{P}}x - \widehat{\boldsymbol{P}}x \right\|_{2}. \tag{1}$$

Average Distance for objects with Indistinguishable views (ADI) [22] is similar to ADD-metric, but adapted for components with symmetric rotation shape and is defined as – utilized for test components (2) and (4).

$$e_{ADI}(\widehat{\boldsymbol{P}},\overline{\boldsymbol{P}};\mathcal{M}) = \underset{i\in\mathcal{M}}{avg}\ \underset{j\in\mathcal{M}}{min} \left\| \overline{\boldsymbol{P}}i - \widehat{\boldsymbol{P}}j \right\|_{2}. \tag{2}$$

Visible Surface Discrepancy (VSD) [23] is proposed to deal with cases of pose ambiguity. VSD is suitable for both symmetric as well as non-symmetric objects and defined as

$$e_{VSD}(\widehat{\boldsymbol{P}},\overline{\boldsymbol{P}};\mathcal{M},I,\delta,\tau) = \underset{p\in\widehat{V}\cup\overline{V}}{avg} c(p,\widehat{D},\overline{D},\tau) \tag{3}$$

$\widehat{V}$ and $\overline{V}$ represent 2D masks of the visible surface rendered from $\mathcal{M}$ ($\widehat{\mathcal{M}} = \widehat{\mathbf{P}}\mathcal{M}$ and $\overline{\mathcal{M}} = \overline{\mathbf{P}}\mathcal{M}$) with estimated pose $\widehat{P}$ and ground truth $\overline{P}$. The tolerance δ is defined to determine visibility.

With distance images $\widehat{D}$ and $\overline{D}$ rendered at the estimated and ground truth, matching cost c is calculated as

$$c(p, \widehat{D}, \overline{D}, \tau) = \begin{cases} d/\tau & if\ p \in \widehat{V} \cap \overline{V} \wedge d < \tau \\ 1 & otherwise, \end{cases} \tag{4}$$

with $d = |\widehat{D}(p) - \overline{D}(p)|$ as distance between the surfaces of $\widehat{\mathcal{M}}$ and $\overline{\mathcal{M}}$ at pixel p. Thereby, τ denotes the misalignment tolerance limiting the allowed range of d.

5. Results and Discussion

Following, the results regarding time efficiency and accuracy for data annotation utilizing the proposed AV-based approach in comparison to the established AprilTag marker technique are presented in Figure 7.

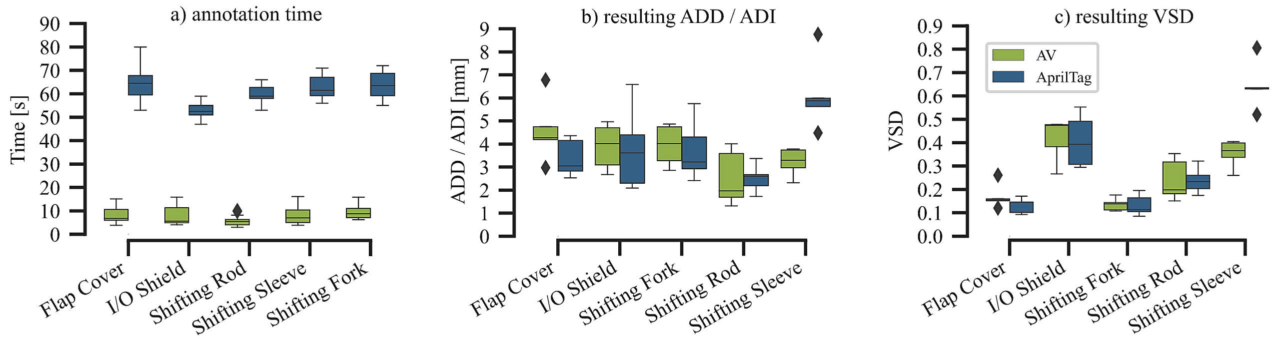


Figure 7: AV annotation versus AprilTag regarding a) the timespan required to create a single annotation, b) the resulting ADD (components: (1), (3) & (5)) and ADI ((2) & (4)) measures and c) the resulting VSD vs. ground truth

5.1 Analysis of the Annotation Time

The experiments show that fiducial marker annotation requires about one minute (cf. Figure 7 (a)). Variance in timespan is mainly caused by the subject experience and the component complexity. The main portion of time consumption is caused by the procedure of marker usage. Additionally, there is geometry dependent preparation time, which increases the required timespan, but scales with annotation numbers. On the other hand, timespans required for AV annotation are 85 % less. On average annotating a component requires about eight seconds as there are no preparation procedures required as well as the faster, immersive interaction with the HMI. Part complexity does not take an equally important role as it does with AprilTag labeling. However, subjects and their experience with using the HMI do have a slight impact.

5.2 Analysis of the Annotation Accuracy

To evaluate the accuracy of the pose estimation of AV versus AprilTag annotations, ground truth is obtained by duplicating the real-world scene and placing models with known poses at measured locations within the scene. Applying the described 6DoF pose evaluation metrics, resulting annotation quality can be quantified.

Regarding ADD or ADI, as applicable, quite similar results are achieved with each annotation method (cf. Figure 7 (b)). The proposed approach is nearly as accurate as annotating with AprilTags, however, requires only about 12 % of the time. For shifting sleeve, which is a hollow part, creating pose labels with AprilTags is more imprecise since measurement of component geometry and center of gravity is complicated. Results for VSD are nearly identical for both approaches (c). With respect to the VSD metric, resulting values below 0.5 are considered as good. Here, AprilTag is outperformed for 6DoF poses obtained for shifting sleeve.

5.3 Demonstration of Operability using Annotated Data under Industrial Application Conditions

The shifting fork component serves as chosen example to demonstrate operability of the annotation method to generate different annotation data types required for an mixed image processing pipeline in a single step (e. g. YOLO + FPN). Final datasets were automatically complemented through varying object and camera poses (see 3.2.2). In the end, 150 component specific measurements are taken into account for YOLO as well as 700 for FPN. Thereby, single- and multi-object setups are evaluated separately. Utilized and evaluated is an initial, unrefined dataset, generated for the system ramp-up stage.

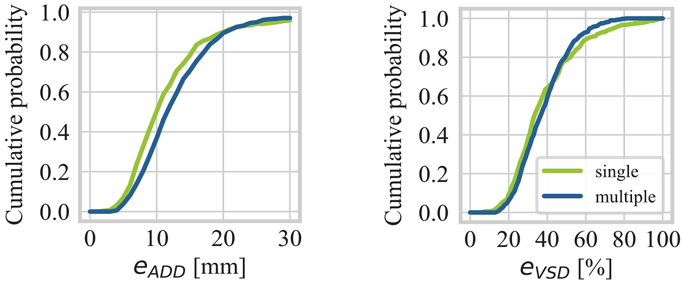


Figure 8: Pose estimation errors retrieved for scenarios with single and multiple objects, visualized as the empirical cumulative distribution function in terms of ADD and VSD metric for shifting fork RGBD real-world dataset

The performance of YOLO in terms of mean average precision (mAP@50IOU) reaches 91.42 % for all five components. Specifically, for shifting fork, an average precision (AP@50IOU) of 95.8 % is obtained. For ADD, a pose estimate is considered correct if error $e_{\text{ADD}} \leq k \cdot d$. Thereby, k is a constant to be chosen and d the largest distance between any pair of model vertices, i. e. the diameter or length. Shifting fork features a dimension of 167 mm × 130 mm × 28 mm (l × w × h). For VSD, a pose estimate is considered correct if $e_{\text{VSD}} \leq 0.5$ is met. Within this evaluation, parameters are set to $k = 0.1$ as well as to $\tau = 30$ mm and $\delta = 30$ mm. In single object scenario a $\text{recall}_{\text{ADD}}$ of 85.3 % and a $\text{recall}_{\text{VSD}}$ of 78.5 % are obtained. In multi-object scenario, a $\text{recall}_{\text{ADD}}$ of 78.7 % and a $\text{recall}_{\text{VSD}}$ of 81.2 % are achieved. e_{ADD} and e_{VSD} visualized as empirical cumulative distribution function are shown in Figure 8. In conclusion, the results obtained are rated as good, especially since solely real-world data without any additional mixed synthetic data is used.

5.4 Analysis of Fault Clearance Potential

To determine the advantage generated by the intervention strategy, a small user study is performed. Therefore, a bin-picking failure is simulated and the timespan resolving the issue is recorded. On average it took about 55 seconds to return the robot into autonomous operation. Compared to manual fault clearance, where the operator walks to the robot, triggers emergency safety circles before taking action, the AV approach saves time. This applies even more for supervising multiple bin-picking systems, since no transit time is required. Besides teleoperated solving, our approach provides direct target pose specification and subsequent automated object handling. Simultaneously, annotated data for dataset refinement is generated.

6. Conclusion and Outlook

In this work, an annotation and refinement method for RGBD data in rigid objects industrial bin-picking is described. The results obtained by the proposed method show a noticeable reduction in effort, measured by the required timespan for annotation, while maintaining high annotation accuracy at least competitive to the state of the art. It provides support for multi-object- and automated annotation as well as different I/O-formats. In addition to pure remote intervention, the HuITL approach enables optimization and enhanced adaptation during system ramp-up through dataset refinement. Although, the method focusses on online data processing, the method is also suitable for dataset generation based on once recorded offline sensor data. Future research will concentrate on hybrid reinforcement- / imitation learning for component grasping.

References

[1] Obermeier, B., Treugut, L. Learning Systems in Hostile Environments, in: , Lernende Systeme 2019, Munich.

[2] Jo, H.-J., Min, C.-H., Song, J.-B., 2018. Bin Picking System using Object Recognition based on Automated Synthetic Dataset Generation, in: 2018 15th International Conference on Ubiquitous Robots (UR). 2018 15th International Conference on Ubiquitous Robots (UR), Honolulu, HI. IEEE, pp. 886–890.

[3] Blank, A., Berg, J., Zikeli, G.L., Lu, S., Sommer, O., Reinhart, G., Franke, J., 2020. Intervention strategy for autonomous mobile robots. wt Werkstattstechnik online 110 (09), 613–618.

[4] Blank, A., Hiller, M., Zhang, S., Metzner, M., Lieret, M., Thielecke, J., Franke, J., 2019. 6DoF Pose-Estimation Pipeline for Texture-less Industrial Components in Bin Picking Applications, in: , IEEE ECMR, Prague, Czech.

[5] Wang, J., Olson, E. AprilTag 2: Efficient & robust fiducial detection, in: , 2016 IEEE IROS 2016, pp. 4193–4198.

[6] Zimmer, W., Rangesh, A., Trivedi, M., 2019. 3D BAT: A Semi-Automatic, Web-based 3D Annotation Toolbox for Full-Surround, Multi-Modal Data Streams, in: 2019 IEEE Intelligent Vehicles Symposium (IV). 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France. IEEE, pp. 1816–1821.

[7] Sager, C., Zschech, P., Kühl, N., 2021. labelCloud: A Lightweight Domain-Independent Labeling Tool for 3D Object Detection in Point Clouds, in: CAD'21 Conference.

[8] Arief, H.A., Arief, M., Zhang, G., Liu, Z., Bhat, M., Indahl, U.G., Tveite, H., Zhao, D., 2020. SAnE: Smart Annotation and Evaluation Tools for Point Cloud Data. IEEE Access 8, 131848–131858.

[9] Marion, P., Florence, P.R., Manuelli, L., Tedrake, R., 2018. Label Fusion: A Pipeline for Generating Ground Truth Labels for Real RGBD Data of Cluttered Scenes, in: 2018 IEEE ICRA. 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD. IEEE, pp. 3235–3242.

[10]Ramirez, P.Z., Paternesi, C., Luigi, L. de, Lella, L., Gregorio, D. de, Di Stefano, L., 2020. Shooting Labels: 3D Semantic Labeling by Virtual Reality, in: 2020 IEEE AIVR. 2020 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR), Utrecht, Netherlands. IEEE, pp. 99–106.

[11]Blank, A., Baier, L., Kedilioglu, O., Zhu, X., Metzner, M., Franke, J., 2021. Efficient AI Adaption using Synthetic Data. wt Werkstattstechnik online 111 (10), 759–762.

[12]Sankaranarayanan, S., Balaji, Y., Jain, A., Lim, S.N., Chellappa, R., 2018. Learning from Synthetic Data: Addressing Domain Shift for Semantic Segmentation, in: 2018 IEEE CVPR. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA. IEEE, pp. 3752–3761.

[13]Budd, S., Robinson, E., 2019. A Survey on Active Learning and Human-in-the-Loop for Medical Image Analysis.

[14]Maadi, M., Aickelin, U., 2021. A Review on Human-AI Interaction in Machine Learning and Insights for Medical Applications. International journal of environmental research and public health 18 (4).

[15]Yu, F., Seff, A., Zhang, Y., Song, S., Funkhouser, T., Xiao, J., 2015. LSUN: Construction of a Large-scale Image Dataset using Deep Learning with Humans in the Loop.

[16]Blank, A., Kosar, E., Karlidag, E., Guo, Q., Kohn, S., Sommer, O., Walter, J., Herbert, M., Sessner, J., Querfurth, F., Metzner, M., Franke, J., 2021. Hybrid Environment Reconstruction Improving User Experience and Workload in Augmented Virtuality Teleoperation. Procedia Manufacturing 55, 40–47.

[17]Kohn, S., Blank, A., Puljiz, D., Hein, B., Franke, J., 2018. Towards a Real-Time Environment Reconstruction for VR-Based Teleoperation Through Model Segmentation, in: 2018 IEEE/RSJ IROS, Madrid, pp. 1–9.

[18]Blank, A., Karlidag, E., Zikeli, L., Metzner, M., Franke, J., 2021. Adaptive Motion Control Middleware for Teleoperation Based on Pose Tracking and Trajectory Planning, in: , Annals of Scientific Society for Assembly, Handling and Robotics, vol. 55. Springer, Berlin, Heidelberg, pp. 153–164.

[19]Heuss, L., Blank, A., Dengler, S., Zikeli, G.L., Reinhart, G., Franke, J., 2019. Modular Robot Software Framework for the Intelligent and Flexible Composition of Its Skills, in: , Advances in Production Management Systems, vol. 566. Springer International Publishing, Cham, pp. 248–256.

[20]Redmon, J., Farhadi, A., 2018. YOLOv3: An Incremental Improvement. Computing Research Repository (CoRR).

[21]Qi, C.R., Liu, W., Su, H., Guibas, L.J., 2018. Frustum PointNets for 3D Object Detection from RGB-D Data, in: 2018 IEEE/CVF CVPR, Salt Lake City, USA, pp. 918–927.

[22]Hinterstoisser, S., Lepetit, V., Ilic, S., Holzer, S., Bradski, G., Konolige, K., Navab, N., 2013. Model Based Training, Detection and Pose Estimation of Texture-Less 3D Objects in Heavily Cluttered Scenes, in: , Computer Vision – ACCV 2012, vol. 7724. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 548–562.

[23]Hodaň, T., Matas, J., Obdržálek, Š., 2016. On Evaluation of 6D Object Pose Estimation, in: Computer Vision ‐ ECCV 2016 Workshops. Springer International Publishing, Cham, pp. 606 ‐ 619.

Biography

Andreas Blank has been a research associate at the Institute for Factory Automation and Production Systems (FAPS) and head of the Industrial and Service Robotics technology field (2014-2021). As PhD student his research is focused on immersive teleoperation, autonomous robot capabilities and human-to-robot skill transfer. He is Production Manager at BIG within Simba-Dickie-Group since 2021 leading factory automation and lean production integration.

Lukas Baier has been a research associate at the Technische Hochschule Ingolstadt (THI) and head of the Production Control and Intralogistics technology field at FAPS (2016-2021). As PhD student at FAPS his research is focused on sensor technologies and data processing for label-free goods identification in intralogistics. He is Senior Expert Operational Excellence at SIEMENS since 2021.

Maximilian Zwingel is a research associate at the Institute for Factory Automation and Production Systems at the FAU Erlangen-Nuremberg with prior research at the THI since 2018. His research is focussed on autonomous mobile robots in intralogistics environments, especially their sensors and navigation.

Prof. Dr.-Ing. Jörg Franke is Head of the Institute for Factory Automation and Production Systems (FAPS) at the FAU. His research focusses on innovative manufacturing processes for mechatronic products. He is involved in leading functions in scientific organisations such as IEEE, CIRP, WGP and 3D-MID. Before his professorship, he led different management positions in industry e. g. at McKinsey & Co., Robert Bosch GmbH, ZF Lenksysteme GmbH and Schaeffler AG.