CPSL

Proceedings of the
Conference on Production Systems and Logistics

CPSL 2021

10^{th} August 2021 – 11^{th} August 2021
Digital Event

by
publish-Ing.

Partners of the Conference

TIB Leibniz-Informationszentrum Technik und Naturwissenschaften Universitätsbibliothek

WGP Wissenschaftliche Gesellschaft für Produktionstechnik

IFA Institut für Fabrikanlagen und Logistik

WZL | RWTH Aachen University

Universität Stuttgart Institut für Industrielle Fertigung und Fabrikbetrieb IFF

fir an der RWTH Aachen

Institut für Werkzeugmaschinen und Fertigungstechnik IWF

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

LPS Lehrstuhl für Produktionssysteme

www.cpsl-conference.com

CPSL

Conference on Production Systems and Logistics

International Conference, CPSL 2021,
Digital Event hosted via publish-Ing.
August 10-11, 2021
Proceedings

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

ISSN: 2701-6277
DOI: https://doi.org/10.15488/11229

publish-Ing.

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 2021), 2021, publish-Ing., S. XX-XX.

Foreword

It is our great pleasure to publish the proceedings of the 2nd Conference on Production Systems and Logistics (CPSL), which was held in an innovative digital concept due to the ongoing COVID-19 pandemic instead of on-site at our partner the University of British Columbia in Vancouver, Canada.

The aim of the CPSL is to provide a platform for scientists from all over the world to present the latest research results, build networks and discuss future research and practice through an annual international conference in the field of production systems and logistics. The highest value is placed on modern, digital and transparent processes for publishing scientific papers, a high-quality review process that all participants can help to shape, and an open access policy so that all findings can be made freely available to the entire scientific community. With this consequent approach of providing added value for science and scientists, we have already been able to win 7 of the most important research institutes for production systems and logistics in Germany as partners and supporters.

Already the first CPSL 2020 in Stellenbosch, South Africa was strongly influenced by the global COVID-19 pandemic. Nevertheless, we managed to hold an unforgettable event with a significantly reduced number of participants on site. The enormous support we received afterwards has given us additional motivation to continue our joint journey with the CPSL and to further improve the concept. Due to the pandemic, the CPSL 2021 could not take place on site at the University of British Columbia as planned, but had to be redesigned as a digital event.
One of our biggest challenges with this format was to make it engaging and fun through an innovative concept. All authors were asked to upload a 10-15 minute video of their paper presentation, which was made available to all other participants one week before the conference. At the conference, all papers were presented in short 3 minute pitches before we started a joint discussion on each paper. This made the conference very entertaining with a high proportion of joint discussions and contributions from the participants.

We would like to express deepest appreciation and gratitude to our sponsors for making all of this possible: publish-Ing., Wissenschaftliche Gesellschaft für Produktionstechnik (WGP), Leibniz Information Center for Science & Technology (TIB), Institute of Production Systems and Logistics (IFA), Werkzeugmaschinenlabor der RWTH Aachen (WZL), FIR an der RWTH Aachen, Institute of Industrial Manufacturing and Management (IFF), Institute of Machine Tools and Production Technology (IWF), Lehrstuhl für Produktionssysteme (LPS) and Institute for Assembly Technology (MATCH). A special thanks goes to all the reviewers for engaging in this demanding but also encouraging review process. Last but not least, we want to thank all of our participants and fellow researchers for sharing their findings, for highly inspiring discussions and for supporting our joint journey.

We are looking forward to meeting you again at the CPSL 2022 in Vancouver, Canada.

Dr.-Ing. David Herberger	M. Sc. Marco Hübner
Conference Chair	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 systems and logistics.

For the submission of a paper, an abstract of approx. 300 words had to be uploaded to the CPSL platform. Abstracts on the following main topics were considered:

- 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 notified and invited to upload a Full Paper. Full Papers had to be no longer than 10 pages (justified exceptions were partially accepted) and had to adhere to a specific template and format provided on the CPSL website. Thereby, papers must meet the standards of good scientific work and aim to share knowledge and relevant findings with the professional community. The authors of the papers are solely responsible for the content of the papers and the mission to provide scientific added value. To ensure high scientific quality and added value, the authors are supported by selected reviewers.

Accordingly, 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. The following key questions, among others, were considered:

- Does the title reflect the contents 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 adapt their papers and submit a revised paper, which were reviewed again in selected cases before final approval.

CPSL

Acknowledgements

Our sincere thanks goes to our outstanding supporters who made this great and interesting conference possible.

Publishing

publish-Ing.

TIB LEIBNIZ-INFORMATIONSZENTRUM TECHNIK UND NATURWISSENSCHAFTEN UNIVERSITÄTSBIBLIOTHEK

Reviews & Scientific Support

WGP Wissenschaftliche Gesellschaft für Produktionstechnik

IFA Institut für Fabrikanlagen und Logistik

WZL | RWTH AACHEN UNIVERSITY

Institut für Werkzeugmaschinen und Fertigungstechnik IWF

Universität Stuttgart Institut für Industrielle Fertigung und Fabrikbetrieb IFF

fir an der RWTH Aachen

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

LPS LEHRSTUHL FÜR PRODUKTIONSSYSTEME

Acknowledgements

A special thanks goes to our outstanding Key-Note speakers, not only for their inspiring and highly interesting presentations but also for their input and contributions in the discussions and Q&A sessions during the conference:

KEYNOTE SPEAKERS 2021

Prof. Dr.-Ing. Volker Stich

Managing Director of FIR at RWTH Aachen & Director of Cluster Smart Logistic

The data driven enterprise – How to close the Gap between today`s and tomorrow`s business

Digitalization is one of the most important topics for each company in future. Data is becoming a core business asset and must be included as a critical resource in any strategic consideration.
The presentation will show, what companies have to do now in order to prepare and develop in the direction of a data driven enterprise, what hurdles must be overcome and if this digitalization-journey is the same for every company.

Prof. Dr.-Ing. Bernd Kuhlenkötter

Director Chair of Production Systems (LPS) and Director Center of the Engineering of Smart Product Service Systems, Ruhr University Bochum

Challenges in Future Production

The keynote will give an overview of technological developments and how they will effect the management and organization of production processes. For example robotic based automation, mobile robotics, additive manufacturing and 5G and their influences on the production of the future will be addressed.

Prof. Dr.-Ing. Thomas Bauernhansl

Director Institute of Industrial Manufacturing and Management (IFF), University of Stuttgart

The Biological Transformation – How the Convergence of Biology & Technology opens up a New Innovation Space for Sustainable Manufacturing

The biological transformation of value creation is the systematic application of knowledge about biological processes with the aim of enabling technology-based sustainable production systems.
Through the transdisciplinary convergence of technologies from the life sciences, engineering and materials sciences, this creates a new innovation space based on so-called biointelligent system architectures.

Table of contents

CPSL

Table of contents

CPSL

Table of contents

CPSL

Table of contents

Table of contents

Table of contents

CPSL

Table of contents

CPSL

Table of contents

2nd Conference on Production Systems and Logistics

A Strategic AI Procedure Model For Implementing Artificial Intelligence

Tim B. Walter[1], Martin Bremer[1], Lars Kaminski[1], Volker Stich[1]
[1]*FIR e.V. an der RWTH Aachen, Aachen, Germany*

Abstract

For most industries, Artificial Intelligence (AI) holds substantial potentials. In the last decades, the extent of data created worldwide is exponentially increasing, and this trend is likely to continue. However, despite the prospects, many companies are not yet using AI at all or not generating added value. Often, an AI project does not exceed its pilot phase and is not scaled up. The problems to create value from AI applications in companies are manifold, especially since AI itself is diverse and there is no 'one size fits all' approach. One often stated obstacle, why many AI projects fail, is a missing AI strategy. This leads to isolated solutions, which do not consider synergies, scalability and seldom result in added value for the company. To create a company-specific AI strategy with a top-down approach, a generic but holistic framework is needed. This paper proposes a strategic AI procedure model that enables companies to define a specific AI strategy for successfully implementing AI solutions. In addition, we demonstrate in this paper how we apply the introduced strategic AI procedure model on an AI-based flexible monitoring and regulation system for power distribution grid operators in the context of an ongoing research project.

Keywords

Artificial Intelligence; Strategy; Framework; Procedure Model; Digital Transformation

1. Introduction

For most companies, regardless of their industry or size, the utilization of Artificial Intelligence (AI) can generate meaningful value for companies [1–3]. Studies predict that AI will be responsible for a third of the German economic growth of the manufacturing industry [3]. Moreover, AI utilization will be necessary to keep pace with global competitors to defend its market position or extend it [4,5]. Thus, the added value that companies can achieve not only consists of a financial dimension but can also include others like competitiveness, better services for customers, or more sustainability. Within the next eight years, Germany's GDP is predicted to increase by 11.3% and its companies' productivity by 4.6% [6]. Within the next five years, a third of the growth is predicted to derive from AI applications [3]. Taking this into account, AI is an useful and necessary field of action for any company as the usage of AI is predicted to be essential to stay globally competitive and thrive economically [4,5]. Furthermore, studies indicate that AI derives meaningful value by increasing companies' revenue and reducing their costs [1]. Moreover, harnessing AI has additional objectives, such as resource deployment minimization, innovation, efficiency increase, and optimization of a company's offer [4].

In principle, companies are open to AI: 46% of German companies are concerned with the issue [4]. AI is perceived as relevant for all companies regardless of their industry and size [2].

DOI: https://doi.org/10.15488/11260 publish-Ing.

Despite the potentials, there are plenty of challenges and pitfalls that hinder the successful implementation of AI applications [4,2]. Although most companies expect new opportunities through AI, nearly half stated that significant investments did not yet add value [7,8]. There are many reasons why an investment in AI applications does not lead to the desired business gains and values. Companies often miss competencies and expertise within their ranks, and their recruitment. Trainings are also obstacles [9,1,4,2,10]. Furthermore, companies lack an AI or data infrastructure or invest in it without a clear understanding of applications and use cases, which leads to unfitting data governance data protection or data strategy [1,2].

Regarding the phrase 'garbage in – garbage out', data quality is an often-underestimated issue that leads to unsatisfying results [9,1]. Moreover, companies often develop isolated pilot solutions without linking the overall strategy if such a strategy exists and do not consider the solutions' scalability [2]. Another obstacle is the missing collaboration across functions, partly due to the missing commitment of the top management and missing acceptance of both employees and customers [1,4,10]. In addition, high investment costs at the beginning, missing best practices, and data privacy and security hinder AI projects. The subject is highly complex – there is no 'one size fits all' solution [4]. For a successful AI application implementation, companies must bring together their technological, cultural, and political domain and prepare the right infrastructure with the right data and talents [2,8].

Some obstacles are not isolated but interrelated with other ones. For example, studies indicate that the lack of an AI strategy contributes to the failure of AI projects and strategic considerations to be vital for a successful implementation of AI [5,8]. Research has shown that a missing AI strategy is one reason for the failure of AI projects and that successful companies have one [9,11,1,4,5,15,16,8]. Furthermore, despite the intensification of research on AI in a business context, aggregated knowledge on this topic is limited, and managers are left with little academic support for implementing AI applications within their companies [8]. A holistic approach for AI implementation, based on an AI strategy, tackles many of the obstacles mentioned earlier and thus enhances the chance of successful value generation [9,11,12,2,13,7,14].

Due to a missing AI strategy, the technology is seldom incorporated into the organization and does not create value. This often leads to an isolated solution that cannot be scaled up comprehensively or solutions that do not fit into the company's strategic direction and contributes little towards company goals [5].

Having an AI strategy would, among others:

1. Improve a company's situation by understanding whether the particular use case is linked to the overall objectives or their organization.
2. Estimate the use case's added value.
3. Prevent projects to remain in the pilot phase by planning their scalability from the start.
4. Define requirements for an AI infrastructure for the entire company, which may plan to implement more than one use case.
5. Consider strategic topics, e.g., legal, privacy, security topics, from the beginning.
6. Ease the management and employees' concerns by communicating the goals and showing them to achieve added value.

The research presented in this paper addresses the aforementioned issues by suggesting the application of a holistic, top-down AI strategic procedure model. In the following, we will present such a procedure model framework. It will enable companies to approach AI projects with a holistic concept, reducing the risk of an AI project failure and supporting its competitiveness and profit.

In the beginning, we will focus on the research gap concerning this subject. We will also define the term 'corporate strategy' that is used in the following. Thereupon, we present a framework for a holistic strategic AI procedure model. Afterward, we apply the suggested framework on an AI-based flexible monitoring and regulation system for power distribution grid operators. Finally, we take an outlook on developments to come.

2. Methodology

We based the development of the proposed AI strategy procedure framework on the method described afterward. First, we conducted extensive desk research to gather information on the current state of the art regarding available AI frameworks and procedure models. Following the collection, we compared the existing strategic AI frameworks and procedure models to gain important and successful factors that affect AI implementation. Using these insights, we derived a procedure considering the working aspects of existing solutions but specifically addressing identified shortcomings. To test the applicability and evaluate the framework, we applied it to a grid operator use case. Finally, to enhance the proposed procedure, we used the results and feedback to further develop the framework.

3. Research results

In the following we will present the results of our research.

3.1 Research gap for AI strategies

Studies show that companies that successfully use AI applications often have an AI strategy with a clear enterprise-level roadmap of use cases that aligns with the corporate strategy [1,16]. Although numerous publications in an industrial context state the need for a holistic AI strategy, there is little scientific research concerning this topic. This might be caused by AI technologies' diverse and non-uniform nature and the strong focus on technical research rather than business-strategic research. Thus, it is necessary to provide academic support for managers implementing AI applications in their companies to reduce the risk of project failure and unwanted results [8].

The strategy has to enable companies to make strategy-oriented AI decisions rather than opportunistic or tactical ones [15]. Moreover, it has to bring together the technological, political, and cultural domains, including data and security issues from the very beginning [17,8]. Unfortunately, research shows that such AI strategies cannot be uniform step-by-step manuals. They rather have to be a framework that allows companies to formulate individual strategies [7].

3.2 Corporate strategy & AI strategy

To be able to define an AI strategy, we first define a corporate strategy. According to Gleißner and Hungenberg, a corporate strategy consists of five components [18,19]:

1. Vision, mission, and long-term goals: A vision describes the long-term target state, which the corporation wants to achieve. Based on this, the mission substantiates three sub-aspects for the company's orientation, namely the field of activity, competence, and values of the company. Out of the mission, long-term company goals are conducted [19].
2. Core competence: The core competencies include those abilities of a company that is essential to operate successfully [18].
3. Business fields and competitive advantages: Business fields describe the field of activity in which a company operates. The market attractiveness and the competitive advantages are its properties as well as the target groups or customers. Out of the customers' needs, the company can deduct products and services [18].

4. Design of the value chain: The value chain is a business process in which value is progressively added to the product. Due to limited resources, the value chain must be designed based on core competencies and competitive advantages [18].
5. Strategic thrust: The strategic thrust consists of factors that may affect the corporation's value. There are three general directions as strategies' main variants: growth strategies, profitability-oriented strategies, and risk-oriented strategies [18].

We define the AI strategy as a subset of the corporate strategy. It comprises 'business fields and competitive advantages' and the 'design of the value chain'. This is due to the four fields of AI application. These are:

1. Internal optimization [20,21],
2. supplementing the existing business area [22],
3. new business areas [22], and
4. digital business models [20].

Except for the internal optimization, all application fields concern the corporate strategy's subfield business fields and competitive advantage. The internal optimization concerns the design of the value chain.

3.3 Applying a top-down-approach for the strategic AI procedure model framework

Several reasons speak in favor of using a top-down approach for an AI strategy. First, it enables coordination throughout the company, which prevents the isolation of AI use cases and promotes synergies [11]. In addition, the coordination of experiments, implementations, selection of AI technologies and vendors across the business prevents the duplication of effort, the usage of competing methods, and multiple vendors [23]. A top-down approach facilitates companies to include strategic goals and consequences into implementing AI projects' running or planned implementation [24–26,14]. Due to these reasons, we propose a top-down approach for the framework of the strategic AI procedure model presented later in this paper.

4. Description of the framework

The framework of the strategic AI procedure model consists of three levels along the top-down-approach as shown in figure 1:

- the corporate strategy level to set the target,
- the meta-level of archetypal AI use cases to mediate between the corporate strategy and AI infrastructure level,
- and the AI infrastructure level, including design fields.

The result is a defined roadmap with prioritized design fields for the implementation.

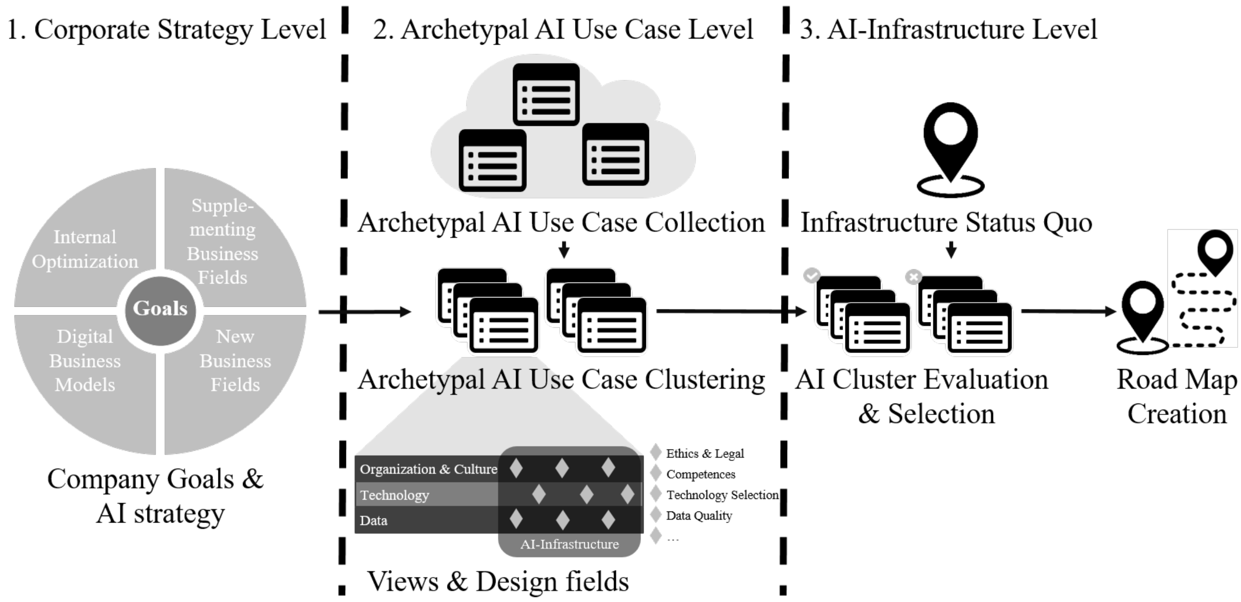


Figure 1: Framework of the strategic AI procedure model

4.1 Corporate strategy level:

Ransbotham et al. have shown that tying a strategy for AI to the company's overall strategy is essential [7]. As stated above, we define an AI strategy as a subset of a corporate strategy, as it comprises the design of the value chain and sometimes the design of the business fields. It follows and aims to realize the corporate goals.

The corporate strategy is a precondition. Based on the enterprise's mission and long-term goals, it provides a frame and extracts specific corporate goals for the AI strategy. The company should consider multiple questions: Which of its core competencies is affected, or does it have to elaborate on new ones? What is the thrust of the AI strategy (is it a growth strategy, profitability-oriented strategy, or risk-oriented strategy)? What is the AI approach, and does it affect business fields (internal optimization, supplementing existing business areas, new business areas, or digital business models)? The long-term goal of implementing AI applications must be pointed out clearly, and it has to fit into the overall corporate strategy.

4.2 Meta level: Archetypal AI Use case collection:

This level contains a collection of archetypal AI use cases. Due to technological progress and wide variety, the collection remains open for additions. This collection shall support identifying use cases under the aspects of identifying synergies with other use cases and technologies, planning scalability, preparing for selection, and supporting understanding of AI capabilities. Supporting this, each archetypal use case contains variables that are important for the selection process. In the following, we list a selection of key questions for several variables, which shall help to determine the relevant variables: *Type of AI:* What is the technology capable of? Does it assess, deduce, or react? Does it imitate human behavior or make rational decisions? *Value creation:* How does the use case create added value? How is it used? What are its limits? *Addressed problems*: Which problems is it tackling? *Input:* What is the required information or data input? *Output:* What is the required information or data output? *Requirements:* What are its requirements on data (amount, quality), domain knowledge, resources, talents, hardware (sensors, GPU, etc.), infrastructure, etc.? *Interconnections:* Are there interconnections with other technologies or use cases, e.g., synergies, dependencies, exclusions, or redundancies? *Interpretability:* Can humans interpret the technology? Do they have to? *Time:* Are there time constraints? How long is the approximate computing time? How long is the approximate training time? *Capacity:* How much server capacity is needed? *Scalability:* Does the use case has to be scaled up? How do we ensure its scalability? *Models:* What models can be used for this use case?

Based on the company goals using the archetypal AI use case collection, use cases can be pre-selected. Important factors are the use cases AI infrastructure requirements, corporate strategy fit, scalability ability, and possible synergies with other use cases or technologies. The single pre-selected AI use cases can then be clustered according to their synergies, value, etc. For instance, an autonomous vehicle use case represents such an AI use case cluster, as it contains several computer vision and decision-making models.

4.3 AI-infrastructure level:

The infrastructure is essential for the success of implementing AI applications. To examine the future AI infrastructure thoroughly, we use three views proposed in the Aachen Digital-Architecture-Management: the **organization** expanded with **culture**, **technology**, and **data** [27]. The AI infrastructure can be divided into three views. For each view, we assign several design fields. A design field contains concrete steps, tasks, and methods to create a part of the AI infrastructure. Each design field belongs to a view, although some are comprehensive and cannot be assigned to only one view. Not necessarily all design fields must be addressed by each company; that depends on the existing infrastructure and the requirements of the to be introduced AI use cases. Examples for the design fields are *Ethics & Legal*, *Cybersecurity (comprehensive)*, *organizational structure*, *roles*, *data governance*, *sourcing & ecosystems (organizational)*, *identification of new technologies*, *platforms*, *user experience (technology)*, *data procurement*, *data storage*, *data processing*, and *data quality (data)*.

On the AI infrastructure level, we propose three steps. First, a status quo analysis needs to take place. It should include infrastructure, system, and data environment analysis. Second, the company should identify relevant design fields for the pre-selected AI use case clusters, specify and compare them to the current infrastructure to estimate the needed effort. Based on this, a value and cost analysis for all clusters is to be conducted. With this information, the company can select its AI applications. Third, the selected AI use case clusters design fields must be customized and prioritized. With the prioritization of all applicable design fields of the selected AI use case cluster, the company can now create a road map for the implementation.

5. Application of the Framework of the Strategic AI Procedure Model

Although the German power grid is one of the most stable grids in the world, measured by minutes of power outages per year, the current energy (higher share of renewable energies) and mobility (battery-powered electric vehicles) transitions and the resulting volatility in the power grid are predicted to harm the grid stability. Higher volatility leads to increased usage and thus wear of the grid components. If grid operators maintained their currently time-based maintenance procedures under these conditions, either increasing power outage times due to grid component-related faults (higher wear and tear) or higher costs due to additional personnel (adjusted maintenance cycles) would be expected. Therefore, grid operators are particularly interested in condition monitoring and predictive maintenance. Developing an AI-based approach for these particular challenges is part of the ongoing research project FLEMING. [28]

While the development of AI algorithms is a fundamental part of the FLEMING project, the project also aims at enabling grid operators to generate value by deploying AI applications. To ensure a strategic approach for implementing this project's solution and for further AI opportunities, the proposed AI strategy procedure framework comes in. We will illustrate this AI Strategy Framework application in this context for a German grid operator who currently has no AI-based solutions in place.

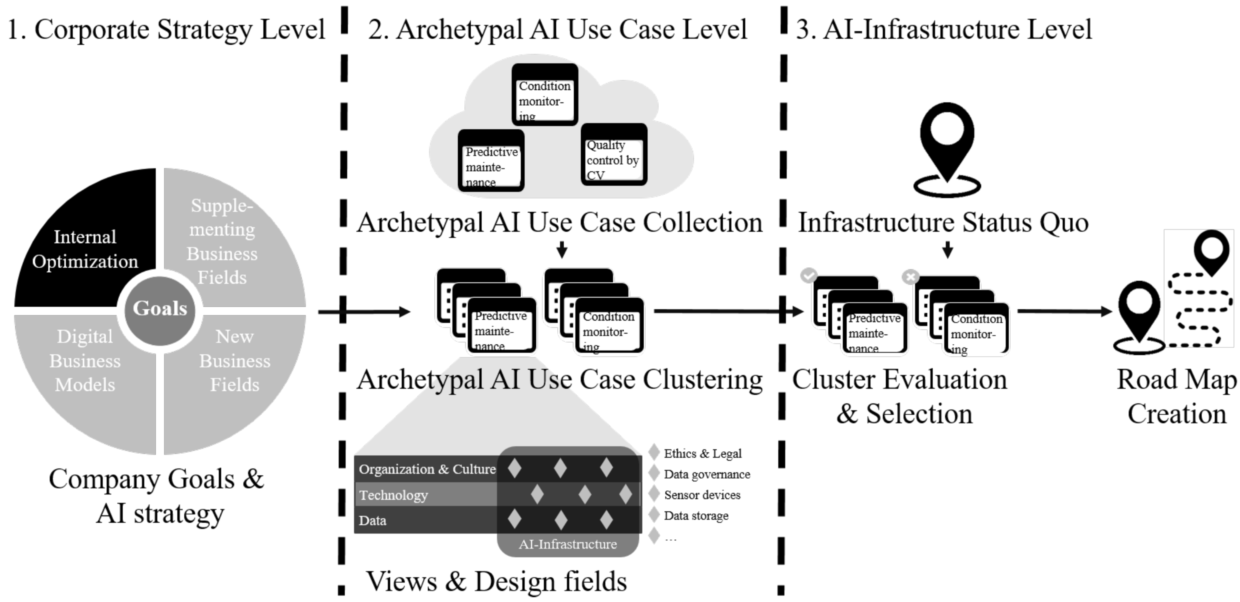


Figure 2: The Framework of the Strategic AI Procedure Model applied for a Grid Operator

As mentioned above, distribution grid operators have the strategic issue of maintaining their service quality, resulting in fewer power outages while being resource efficient. By using AI applications, they hope to strengthen their core competencies in the sense of efficient and reliable services. They focus on a profitability-oriented strategy, which does not negatively affect the corporate strategy. Thus, the AI approach is one of internal optimization. We then matched these company goals to the archetypal use case collection. For internal optimization, the collection proposes archetypal use cases as ‘predictive maintenance’, ‘condition monitoring’ or ‘quality control by computer vision’. These use cases can be considered by themselves or be clustered to a multiple-use case solution, considering their synergies, corporate strategy fit, and scalability ability. Since ‘quality control by computer vision’ does not support its goals, it is rejected.

In contrast, the first two use cases are pre-selected and, if applicable, specified (‘predictive maintenance for switchgears’). Each pre-selected archetypal use case cluster or single-use case contains comprehensive design fields and those connected to a view (organization & culture, technology, data), representing necessary elements of a holistic AI infrastructure. For the use case “predictive maintenance,” some relevant design fields could be, for example, *ethics & legal, cyber-security (comprehensive view), data governance, change management (organization & culture view), sensor-devices, platform infrastructure (technology view), data collection, storage*, and *quality (data view)*. They can be compared to the current infrastructure after conducting corresponding analyses. Moreover, the design fields can be specified and identified, which are necessary for the transformation or building of the novel AI infrastructure. For example, if the grid operator already has all the necessary sensors needed for “predictive maintenance” operating, there is no need to further stress this design field. In the following, the clusters can be evaluated regarding their value and costs, which leads to their selection.

At last, the selected cluster’s design fields are prioritized, and a roadmap for implementing it is developed. Since this approach is being developed in an ongoing research project, no validation exists at this time. However, the results from the research project will be examined and validated in more detail in subsequent publications.

6. Summary & Outlook

This paper briefly presents the opportunities and pitfalls of AI applications for the industry. We identified a missing AI strategy as a major obstacle to a successful AI implementation. To tackle this obstacle, we introduced a framework of a strategic AI procedure model and applied it to a grid operator in the context of an ongoing research project. The framework is the subject of further development. We will create a collection of archetypal use cases and elaborate on relevant factors of the AI use cases. Moreover, we will complete a list of the design fields as far as possible and specify each field. Finally, we desire further research for the selection process of AI use cases and their cost and value analysis.

7. Acknowledgements

The research in this paper is part of the project FLEMING, funded by the German Federal Ministry for Economic Affairs and Energy (project number 03EI6012D).

References

[1] Cam, A., Chui, M., Hall, B., 2019. Global AI Survey: AI proves its worth, but few scale impact, 11 pp.

[2] Hartmann, P., Liebl, A., Waldmann, A., Brakemeier, H., 2019. Applying AI: The elements of a comprehensive AI strategy, 14 pp.

[3] Seifert, I., Bürger, M., Wangler, L., Christmann-Budian, S., Rhode, M., Gabriel, P., Zinke, G., 2018. Potenziale künstlichen Intelligenz im produzierenden Gewerbe in Deutschland, 68 pp.

[4] Gentemann, L., Lange, B., 2020. Germany is Learning AI.: How Companies are Making Use of Digital Technologies, 39 pp.

[5] Heinz, S., 2020. The 6 key elements of an AI strategy.

[6] Kirschniak, C., 2018. Auswirkung und Nutzung von künstlicher Intelligenz in Deutschland, 18 pp.

[7] Ransbotham, S., Khodabandeh, S., Fehling, R., LaFountain, B., Kiron, D., 2019. Winning With AI: Pioneers Combine Strategy, Organizational Behavior, and Technology, 27 pp.

[8] Reim, W., Åström, J., Eriksson, O., 2020. Implementation of Artificial Intelligence (AI): A Roadmap for Business Model Innovation. AI 1 (2), 180–191.

[9] agiles Group, 2020. 3 reasons why many AI projects fail. https://agiles.com/en/blog-3-reasons-why-ai-projects-fail/. Accessed 7 June 2021.

[10] Sallaba, M., 2020. State of AI in the Enterprise: Ergebnisse der Befragung von 200 AI-Experten zu Künstlicher Intelligenz in deutschen Unternehmen, 24 pp.

[11] Boll-Westermann, S., Faisst, W., 2019. Neue Geschäftsmodelle mit Künstlicher Intelligenz: Zielbilder, Fallbeispiele und Gestaltungsoptionen, 76 pp.

[12] Caner, S., Bhatti, F., 2021. A Conceptual Framework on Defining Businesses Strategy for Artificial Intelligence.

[13] Kitsios, F., Kamariotou, M., 2021. Artificial Intelligence and Business Strategy towards Digital Transformation: A Research Agenda. Sustainability (4).

[14] Schoonhoven, J.J., Roelands, M., Brenna, F., 2019. Nach dem Hype: Was Führungskräfte über die erfolgreiche Implementierung künstlicher Intelligenz wissen müssen, 20 pp.

[15] Joshi, A., Wade, M., 2020. The Building Blocks of an AI Strategy: Organizations need to transition from opportunistic and tactical AI decision-making to a more strategic orientation. https://sloanreview.mit.edu/article/the-building-blocks-of-an-ai-strategy/.

[16] Kruhse-Lehtonen, U., Hofmann, D., 2020. How to Define and Execute Your Data and AI Strategy. Harvard Data Science Review.

[17] Leprince-Ringuet, D., 2020. AI for Business: What's going wrong, and how to get it right. Managing AI and ML in the enterprise 2020, 3–9.

[18] Gleißner, W., 2004. Future Value: 12 Module für eine strategische wertorientierte Unternehmensführung.

[19] Hungenberg, H., 2014. Strategisches Management in Unternehmen. Springer Fachmedien Wiesbaden, Wiesbaden, 605 pp.

[20] Gentsch, P., 2019. Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business – Konzepte und Best Practices, 2nd ed. Gabler Verlag.

[21] Gläß, R., 2018. Künstliche Intelligenz im Handel 1 – Überblick: Digitale Komplexität managen und Entscheidungen unterstützen, 1st ed. Springer Vieweg.

[22] Grell, A., 2019. Digitale Transformation bei einem technischen Dienstleister, in: Dahm, M., Thode, S. (Eds.), Strategie und Transformation im digitalen Zeitalter. Inspirationen für Management und Leadership, 1 ed. Springer Fachmedien, Wiesbaden, pp. 309–325.

[23] Ammanath, B., Hupfer, S., Jarvis, D., 2020. Thriving in the era of pervasive AI, 28 pp.

[24] Braadbaart, J., 2019. An illustrated guide to AI for the handsome corporate executive. Towards Data Science, May 17.

[25] Hofmann, P., Jöhnk, J., Protschky, D., Stähle, P., Urbach, N., Buck, C., 2020. KI-Anwendungsfälle zielgerichtet identifizieren. Wirtschaftsinformatik & Management 12, 184–193.

[26] Mazzei, C., Duffy, N., 2017. Putting artificial intelligence (AI) to work: Innovation mates: insights on the latest disruptive technologies.

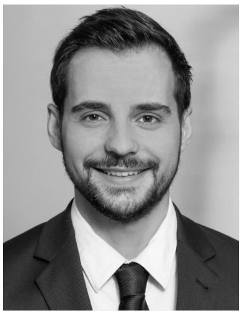
[27] Hicking, J., Wenger, L., Abbas, M., Benning, J., Bremer, M., Clemens, F., Engländer, J., Heimes, P., Henke, L., Kaminski, L., Kremer, S., Niederau, M., Seelmann, V., Stroh, M.-F., Walter, T., 2020. Aachener Digital-Architecture-Management: Wegweiser zum digital vernetzten Unternehmen. FIR e.V. an der RWTH Aachen, Aachen, 32 pp.

[28] Hoffmann, M.W., Wildermuth, S., Gitzel, R., Boyaci, A., Gebhardt, J., Kaul, H., Amihai, I., Forg, B., Suriyah, M., Leibfried, T., Stich, V., Hicking, J., Bremer, M., Kaminski, L., Beverungen, D., zur Heiden, P., Tornede, T., 2020. Integration of Novel Sensors and Machine Learning for Predictive Maintenance in Medium Voltage Switchgear to Enable the Energy and Mobility Revolutions. Sensors (Basel, Switzerland) 20 (7).

Biography

Tim B. Walter (*1993) has been a scientific researcher and project manager of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2020. In his current position as part of the Information Management Division, he supports companies in various industries in the design and implementation of the information logistics of business processes.

Martin Bremer (*1989) started working as a project manager at FIR at RWTH Aachen University in 2017. In his current position within the technical group IT complexity management, he specialized within IT architecture management, the development of transformation strategies, enterprise cloud transformations and Business-IT-Alignment.



Lars Kaminski (*1993) started working as a project manager at FIR at RWTH Aachen University in 2019. In his current position within the technical group IT complexity management, he specialized within IT architecture management and the development of transformation strategies with respect to energy management systems and information security.

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 led the management of European plant logistics. His research focuses on operations management, logistics and business applications systems.

2nd Conference on Production Systems and Logistics

Life Cycle Oriented Planning Of Changeability In Factory Planning Under Uncertainty

Lennart Hingst[1], Yeong-Bae Park[1], Peter Nyhuis[1]
[1]Institute of Production Systems and Logistics (IFA), Leibniz University Hannover, Germany

Abstract

Factory planning and factory operation collectively form a major part of the factory life cycle. The growing awareness of uncertainties throughout the factory life cycle is not only a consequence of recent events, but also the realization that factories are operating in an increasing turbulent environment. In factory planning, various sources of risk (e.g. location, process) can cause uncertainties due to deviations in the planning parameters (e.g. filling quantity load carrier) that affect the capacities to be dimensioned. During factory operation, sources of uncertainty (e.g. lead time, quality) expose the factory to numerous events that may disrupt their business process (e.g. machine failure). Despite these short-term and random risk events, factories are confronted with long-term change drivers (e.g. new product variants) in the course of their life cycle due to continuously rising requirements. Instead of responding reactively in case of uncertainty, it is much more appreciated to proactively prepare the factory for the uncertainties. It is the task of factory planning to gear up the factory for whatever uncertainties may occur over the factory life cycle. But the ability to change goes hand in hand with higher cost levels, either in the form of capital or operational expenditure depending on the type of changeability. Since there is a wide range of factory planning measures that allow the factory to be configured in different ways, factory operation must be considered in order to select a suitable factory type from life cycle perspective. Therefore, the goal of this paper is to integrate an approach in the sense of risk management within factory planning. Consistent factory types for coping with uncertainties are defined in order to present a way on how to position the factory in the area of tension between profitability and changeability.

Keywords

Factory Planning; Uncertainties; Changeability; Factory Life Cycle

1. Introduction

In the course of a turbulent environment, factories are confronted with various uncertainties [1]. Specifically, the manufacturing sector encounters internal and external change drivers arising from so-called megatrends [2], but also short-term and operational risks or disruptions [3]. The effects of the uncertainties are further intensified by the ever-increasing dynamics and synchronization of processes in the factory. Even minor incidents can have significant consequences for factory operations. [4] As part of factory planning, suitable measures must be taken in advance so that possible consequences can be reduced to a minimum in the course of the factory life cycle [5]. The goal is to eliminate or reduce the probability of occurrence. If the effects of uncertainty manifest themselves nevertheless, the factory must at least be able to reduce them quickly. Therefore, an approach must be integrated within factory planning in order to be able to assess the exposure to uncertainties and determine the required degree of changeability of the factory planning variants. The

DOI: https://doi.org/10.15488/11276

literature describes various forms of changeability that can mitigate both short-term and long-term uncertainties. In addition to the proven concepts of flexibility and transformability, resilient or robust factories, among others, are promoted in the wake of current events. A precise comparison of the terms has not taken place yet. Furthermore, factory planners are faced with the question of what level of uncertainty is to be expected over the factory life cycle and what level of changeability seems appropriate from an economic point of view. The goal of this paper is to provide a solid basis for managing uncertainties in the context of factory planning. Potential effects of uncertainties throughout the factory life cycle will be characterized and different concepts of changeability will be described as factory types. By linking different kinds of uncertainties to possible factory types, a first approach is given on how to position the factory in the field of tension between economic efficiency and changeability.

2. Basics and Need for Research

The factory as a socio-technical system is composed of both technical and social elements that have mutual interactions [6]. The factory and all its associated elements operate in a turbulent environment. In addition to the constantly emerging changes in this environment, internal changes also influence the factory and its elements [7]. This requires a wide variety of responses and adjustments at different times, which are usually planned, prepared and implemented as part of factory planning. In the literature, there are numerous definitions of the term factory planning, which have been combined in the VDI 5200 as a "systematic, objective-oriented process for planning a factory, structured into a sequence of phases, each of which is dependent on the preceding phase that extends from the setting of objectives to the start of production“ [8]. In addition to factory planning, the life cycle of a factory consists of realization, ramp-up, factory operation and shut down. The manufacturing of products takes place during factory operation using raw, auxiliary and operating materials [9]. The processes involved form the basis for order fulfillment. In addition, there are the processes of steering and controlling the operations in a factory [1].

A number of approaches dealing with the life cycle and life cycle management of a factory can be found in the existing literature. The overall factory life cycle is composed of different life cycles of the individual elements of a factory, which must be aligned and therefore requires a holistic, end-to-end planning activity [10]. Due to the turbulent environment, factory planning projects are becoming increasingly necessary and are consequently triggered at ever shorter intervals. In the meantime, they have become an interdisciplinary ongoing task for companies. [11] In the course of factory planning, the factory is usually designed for a certain time horizon of up to 10 years, depending on the planning case, and thus only covers a part of the factory life cycle [9]. Uncertainties must be taken into account within this time horizon in order to ensure that the factory is future-proof. To evaluate the effectiveness and efficiency of possible changeability concepts in the context of factory planning, a life-cycle oriented evaluation of the factory over the time horizon under consideration is required. Both initial capital expenditure for realizing the factory planning variants and operational expenditure in the course of factory operation must be included. Methods for evaluating costs over the life cycle are summarized under the term life cycle costs (LCC). [10] Literature reviews in the field of life cycle costs have shown that a quantitative evaluation of a factory has not been conducted yet [12]. The turbulent factory environment was not included in previous reviews though. Therefore, possible approaches for evaluating changeability in a factory have not been examined from a life-cycle perspective yet.

The required level of changeability is determined based on the assessment of uncertainties. The origin in the assessment of uncertainties comes from the concept of risk management, which originates from the insurance industry [13]. Initially, various process models in the literature have dealt with risk management as a systematic and continuous approach [14,15,13,16,17]. Subsequently, the focus shifted towards the assessment of individual, short-term uncertainties, such as fluctuations in demand [18–20], disruptions in the production process [21–24] or supplier or quality issues [25,26], in order to derive suitable measures for dealing with short-term uncertainties. Only a few approaches address the effects of different categories of short-term uncertainties on the entire factory [27–29], also lacking a life-cycle evaluation of cost effects. Same applies to an approach, which not only takes into account short-term uncertainties but also long-term uncertainties such as changes in technologies, laws, products, etc. [30]. Mostly, similar approaches in the literature focus exclusively on long-term uncertainties and attempt to evaluate respective responses either monetarily [31–33], to investigate the interactions between uncertainties and the factory [2,34,35] or to develop process models for managing long-term uncertainties based on the scenario analysis [36,37] or control systems [38–40]. In the light of current developments, new terminologies for the concept of changeability continue to emerge. There is a lack of an approach that combines the different strategies of changeability with the claim to consider both short-term and long-term uncertainties within factory planning. The key requirement is the targeted adaptation of the influenced factory elements to the expected uncertainties during factory planning or operation in a life cycle cost-efficient manner based on the assessed level of impact and probability.

3. Changeability through factory planning from life cycle perspective

The first step in overcoming the aforementioned shortcomings will be to establish a universal understanding of the concept of changeability in the context of factory planning throughout the paper as a basis for further work. For this purpose, a distinction between short-term and long-term uncertainties is made first. This is followed by an explanation of how these uncertainties are taken into account in the factory planning process. Finally, various strategic factory types are described in the context of changeability and both differences and similarities between the factory types are identified.

3.1. Differentiation of short- and long-term uncertainties

In transaction cost theory, uncertainty is associated with the variability of outcomes, lack of knowledge about the distribution of potential outcomes and uncontrollability of outcome attainment [41]. Figure 1 shows that this uncertainty leads to changes that are the result of long-term developments in the factory environment. Alternatively, it emerges directly from the factory as a complex socio-technical system. Complexity drivers in factory planning or operations increase the level of uncertainty leading to an occurrence of events that cause disruptions in factory operations.

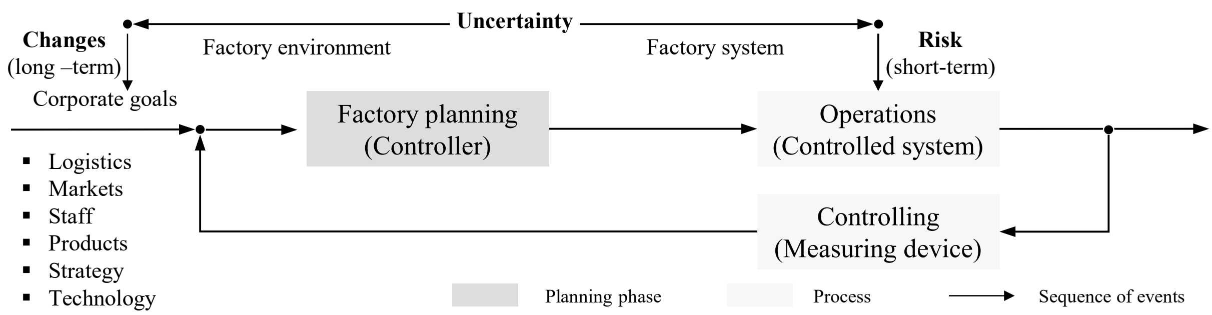


Figure 1: Uncertainty influencing the control system of factory planning, based on [42]

The control system of factory planning provides for an adjustment of the factory by factory planning as soon as controlling detects a deviation of the target/actual values of strategic and operative key figures. The target values can be derived from the corporate goals, while the actual values are recorded via the controlled system. [42] In the following, it will be explained how short- or long-term uncertainties lead to a negative influence of the actual values or an adjustment of the target values.

Uncertainty is present in any complex system. Sources of uncertainty may include lead time, market demand, product quality, and information flow among others [41]. The resulting events are usually random and have a probability of occurrence. They are disruptive, have a relevant impact on performance, and are sometimes difficult to anticipate. Therefore, an uncertain event occurring on a short-term basis is also defined as a risk, whose occurrence will have an impact on the achievement of one or more goals. [43] Regarding the interpretation of the term, a distinction can be made between a cause-related and an effect-related definition of risk. The risk can be considered as a causal factor that triggers an event (e.g. change of supplier in factory planning, faulty operation of equipment in factory operation). According to the effect-related definition, risk is understood as the possible occurrence of an event that negatively affects the achievement of a goal. [44] Consequently, risk is defined as the potential damage of a future event (e.g. change in container filling quantity in factory planning, machine failure in factory operation).

In addition to risks as short-term uncertainties in factory planning and operation, change drivers appear as medium- and long-term uncertainties that lead to ever new requirements for the factory. They result from megatrends that have a global impact in all areas of society and are described as long-term developments with major economic, political and social relevance. [42] Due to the strong pressure of megatrends for change, companies are forced to adapt their strategy and corporate goals. Some prominent examples of current megatrends include climate change and demographic change [45]. They influence the business models of companies and their entrepreneurial actions [46]. However, megatrends only have an indirect influence on the factory. A direct influence results from the caused change drivers such as customer demands, the sales markets or the product and technology life cycle, which describe the effects of megatrends on the environment of manufacturing companies [1]. Constantly new requirements for the factory require increasingly frequent adaptions [2]. Various authors have developed catalogs for change drivers. An exemplary catalog was consolidated as part of the study of relationships between megatrends and change drivers [5]. The catalogs support companies in identifying relevant change drivers in order to respond to a changing environment at an early stage.

3.2. From changeability planning to life cycle oriented consideration of uncertainties in the factory planning process

During factory operation, defined threshold values can be continuously breached or exceeded. In this case, the factory no longer meets the desired requirements, so that a new factory planning process has to begin. Changeability planning facilitates the comparison of actual and target values and thus supports the planning decision to initiate a planning process. As factory planning is a recurring process, monitoring the threshold values to trigger a further planning loop is considered an essential part of changeability planning. [2] The temporal classification in figure 2 is done with the help of the factory life cycle.

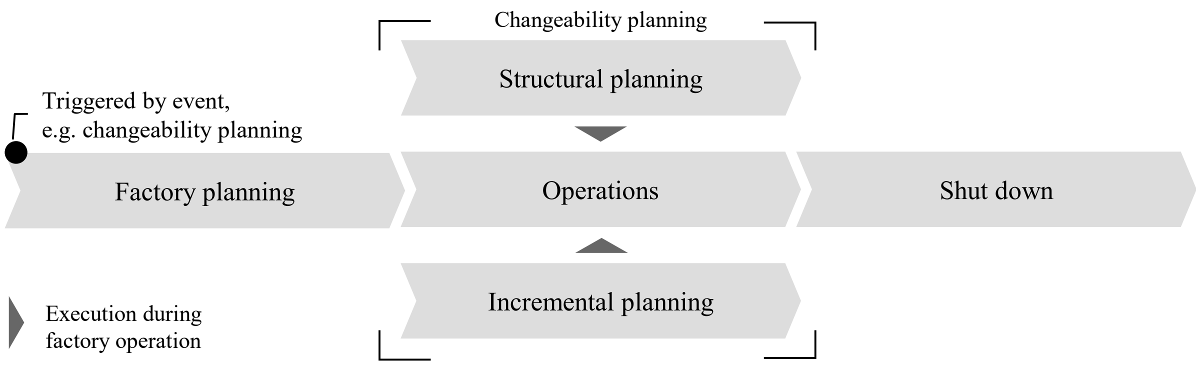


Figure 2: Classification of changeability planning in the factory life cycle, based on [2]

The new planning of a factory is followed by factory operation. Further factory planning loops are triggered by certain events during factory operation. These are captured by changeability planning. Two types of changeability planning can be distinguished. Structural planning involves the ability to recognize trends early on in order to initiate the necessary measures. In contrast, incremental planning involves a high level of responsiveness in order to be able to react to short-term changes in the production environment. [42] As the incremental planning seeks to establish a certain capability for turbulence in the factory, it equates risk management in the following.

At the beginning of the incremental planning, the company defines its risk awareness [47]. To this end, certain risk strategies must be defined for determining which risks are to be taken and how they are to be handled in case of the occurrence of risk events. The individual phases consist of risk identification, risk assessment, risk control and risk monitoring [48]. In risk identification, all relevant risks are recorded holistically. This requires the exposure of all sources of danger, causes of damage or potential for disruption. Risk assessment focuses on the estimation and evaluation of risks. The purpose is to quantify the impact of risks in order to be able to estimate danger potentials. [16] This step is necessary in order to correctly assess and prioritize the need for further action [25]. The cause of risk as the actual reason for the occurrence of risk determines which risks can be influenced by the company. Risks of external origin cannot be influenced, or can only be influenced with great difficulty. Risk control involves the treatment of the previously identified and assessed risks. A distinction is made between risk avoidance, risk reduction, risk diversification, risk transfer and risk acceptance. [48,25,47] The main factors influencing the determination of measures are the risk class, the level of impact, and the probability of occurrence of the risks [47]. Risk monitoring is at the end of incremental planning. It runs alongside the processes and monitors the dynamic development of the risks [25]. Incremental planning of changeability does not necessarily result in the initiation of the factory planning process, as it is limited to an adjustment of the factory in terms of detailed planning [42]. As a result, very limited degrees of freedom are available. The focus of incremental planning especially lies on the mitigation of short-term disruptions.

If the factory cannot be adequately prepared for the uncertainties through incremental planning, the factory is adapted using structural planning of changeability. In the process of questioning organizational and structural relationships in the factory, it is mandatory to go through the factory planning process again [37]. The success of the structural change process is not so much dependent on the ability to react, but rather on the timely recognition and implementation of necessary measures [42]. Emerging trends must be identified at an early stage through forecasting and must be shaped proactively by creating a new solution space. Structural planning consists of the monitoring of changes, the assessment of the quality of changeability and the determination of corrective measures in case an increase of changeability is necessary. It is divided into the two phases monitoring and assessment of changeability. [31,2] The assessment is only triggered if unknown changes in the environment are registered that have an impact on the factory. For this purpose, potential change drivers are derived from megatrends as a first step in monitoring. Relevant developments of the drivers as well as the resulting need for change are concluded. The need for change results from the change dimensions of quantity, variants, costs, time or quality. [2] At the end of the monitoring phase, the processes and elements of the factory that are affected by the need for change must be determined. This is followed by an assessment of the factory to determine whether the changeability of the affected processes and elements is sufficient to cope with possible future developments. If a need for action is identified, corrective measures to increase the changeability must be specified and implemented. [31,2]

Changeability planning can be applied in cycles during factory operation. Another possibility as an addition to the continuous monitoring of uncertainties during factory operation is the use of the planning method in the factory planning process. [2] By identifying future requirements, these can be addressed in conceptual and detailed planning [37]. Provided that the factory's ability to change is identified as an essential target field, possible risks and change drivers that may occur during the factory life cycle should be identified when setting the objectives in factory planning. Depending on the necessary degree of changeability identified in this way, the factory can be designed accordingly. There are various strategies of changeability available to the factory planner in this regard, which can result in different factory planning variants. The final selection of a preferred variant can only be made on the basis of a factory lifecycle-oriented evaluation by anticipating the later factory lifecycle and testing how the various planning variants deal with the uncertainties. This allows the capital and costs expenditures for changeability to be compared with the generated benefits along the factory lifecycle. In the following, the different strategies of changeability will be briefly introduced and distinguished from each other in the form of factory types.

3.3. Selection of factory types in the context of changeability for coping with uncertainties

Change is also defined as the "agreed establishment of a new state instead of the previous state" [49]. The ability to cope with change triggered by uncertainty can involve different levels and parts of the factory. Only isolated sections of the overall system are subject to a change process with first-order change. In contrast, all organizational dimensions are affected in the case of second-order change [50]. From a factory planning perspective, second-order changes are particularly relevant. In order to give an overview of different strategies of changeability in the literature, a preliminary review was performed and the resulting factory types are described below (Table 1).

Table 1: Comparison of the different strategies of changeability as possible factory types

	Robustness	Resilience	Flexibility	Transformability	Agility
Robustness	Robustness represents the insensibility to internal and external disruptive events in relation to the performance of the production system.	Resilience has a dynamic behavior compared to static system property of robustness.[3] Adaptations are allowed (require opex instead of capex), in the same way as temporary performance drops.[21]	Flexibility is also intended for dealing with mid-term uncertainties. In contrast to robustness with fixed capacities, reactive system adaptations are carried out that are reversible.[51]	Transformability is also intended for dealing with mid- and long-term uncertainties. [6] In contrast to robustness, potential is held for specific events and changes are made responsively when needed. [51]	Agility is established for a complete company and network. [52] It does not require pre-planning, thus future developments do not have to be known. [53]
Resilience	Both types are conceived in the short term and operationally. The goal is a certain resistance to disturbances. Resilience can be a kind of balance between robustness and transformability. [45]	Resilience refers to the ability of a system to endure certain disruptive events without failing completely and to return to its original state within a short time after the disruptions have ceased.	Flexibility is also intended for dealing with mid-term uncertainties. Classification according to flexibility types. Besides the swiftness of adaptation, a strong focus is put on simplicity of adaptation. [51]	Transformability is also intended for dealing with mid- and long-term uncertainties. Increased focus on change drivers with the help of specific enablers that require capital expenditures for implementation. [9]	Agility is established for a complete company and network. Very long-term impact horizon. [52,54]
Flexibility	Implementation of predefined fields of action in the form of flexibility corridors [20] and robustness limits to deal with potential developments within this field of action. [53]	Both types are rather operational. Fast reactions are a crucial factor for success. [20] Flexibility is an implicit resilience characteristic. [21]	Flexibility refers to the ability to adapt to change requirements simply in a short time and without major investment within defined flexibility corridors, with no significant changes to the structure.	Flexibility represents reversible changes, affected operating costs, and maintained corridors, [30] whereas transformability represents irreversible changes that go beyond corridors and require an activation effort. [3,9]	Agility refers to a complete company and network. Much more comprehensive, e.g. by switching between entire product families. [55,56]
Transformability	Proactive strategies by means of capital expenditures and thus considered rather irreversible once implemented.[3,53]	Resilience can be achieved through a balance between robustness and transformability. Human creativity is important. [21,45]	Represent the two categories of conventional changeability. [42] Transformability has historically emerged from flexibility.	Transformability is the potential to carry out organizational, technical, spatial and logistical changes outside the flexibility corridors provided at all system levels by means of transformation enablers.	Agility refers to a complete company and network and is thus more comprehensive. [52,9] It does not require preplanning, thus future developments do not have to be known. [53]
Agility	-	Both types have quickness as a characteristic, as well as proactive and reactive elements in common. [21,52]	Both are prone to change/ adaptations. Flexibility is an integral part of agility. [52]	Both require extensive adaptation and are designed to cope with increasing market complexity in a strategic way. [57] Transformability is an integral part of agility. [52]	Agility gives the entire company the ability to respond to uncertainties and impulses for change strategically within the shortest possible time.

Legend: (white) Definition; (light grey) Similarities; (dark grey) Differences

Five different strategies of changeability have been identified in the literature as possible strategic factory types. First, a closer look is taken at robustness, for which divergent definitions exist. However, previous analyses agree that robustness describes the ability of a system to be insensitive to changing environmental influences [3]. To some extent, it is added that system adaptations do not need to be made in order to cope with these influences [51]. In case of changing environmental conditions or deviations caused by disruptions, the function of the system can still be maintained [58]. Thus, a robust production system can withstand a certain level of stress without suffering deterioration or loss of functionality [59].

The word resilience comes from the Latin verb "resilire" meaning "to spring back, bounce back". Interdisciplinary, resilience is described as the ability to handle critical situations, to prevent damage when disruptions occur, and to return to the previous state as quickly as possible through rapid recovery. [60] Transferred to a production system, a resilient system is allowed to leave the steady state for a short period of time. After a brief drop in the performance level following the occurrence of the disruptive event, the system must be able to return to its original performance as quickly as possible through self-regulation [53]. In this regard, humans as intelligent elements of a socio-technical system are at the center of resilient production systems due to their anticipation, interpretation, and decision making [21]. Other central notions include interconnectivity, resistance, adaptivity, decentralization, and learning capability [61].

At the turn of the millennium, the term flexibility was one of the most frequently discussed approaches with regard to the ability of companies to change [9]. Over 50 different definitions and interpretations for the term flexibility have already been identified in 1990 [62]. Meanwhile, the concept of changeability continued to evolve. Former elements of flexibility are assigned to other strategies of changeability today. Essentially, flexibility refers to the ability of a factory to adapt quickly and with very little cost within flexibility corridors that are defined at the point of factory planning [54]. There is an increased focus on the simplicity and reversibility of adaptations [51]. Due to the flexibility corridors, the response options are limited [42]. The previous explanations correspond to the definitions of short- or medium-term as well as static flexibility. The literature also provides definitions of long-term and dynamic flexibility [31,37], part of which is associated with transformability.

Transformability puts the factory into a commitment to change using a transformation potential, whose activation expands the original function or shifts the flexibility corridors [40]. Irreversible changes can be made responsively when needed, which are beyond the flexibility corridors held in reserve [54]. As a result, the time and cost required to prepare the necessary adaptations is significantly higher compared to the concept of flexibility. As soon as the preparations have been completed, organizational, technical and logistical changes can be implemented outside of maintained flexibility corridors in a short time when required, with low investments and taking into account the interactions of the system elements. [37] For example, the infrastructures for change are already configured and in place before specific needs for change are known or arise [9].

The final element is agility, which is defined in broader terms than the conventional ability to change. The strategic focus also includes units outside production, such as sales, purchasing and controlling [9]. An agile approach empowers the company to change its entire production networks or its entire product and service portfolio [54]. This includes measures such as relocating the production site or switching from multiple to single sourcing [31]. By reducing planning activities to a minimum, it is also possible to respond to change drivers immediately during the planning phase. Thus, agile systems are able to cope with unforeseen and unpredictable events [53]. Generally, the identification of market opportunities is addressed through an agile business model, so that a prompt fulfillment of every customer request can be achieved [32].

The factory types derived from the strategies of changeability are overlapping. To some extent, the terms are used mutually in the literature when defining strategies of changeability. One possible definition of transformability, for example, is a combination of the four aspects robustness, flexibility, agility and adaptability [51]. A clear distinction of the strategies of changeability does not exist, partly because the concept has been extended gradually over the last decades. Based on flexibility, transformability and agility were complemented first. Due to the current circumstances, the terms resilience and robustness are currently used repeatedly in the literature. The conclusions of this paper have been summarized in Figure 3.

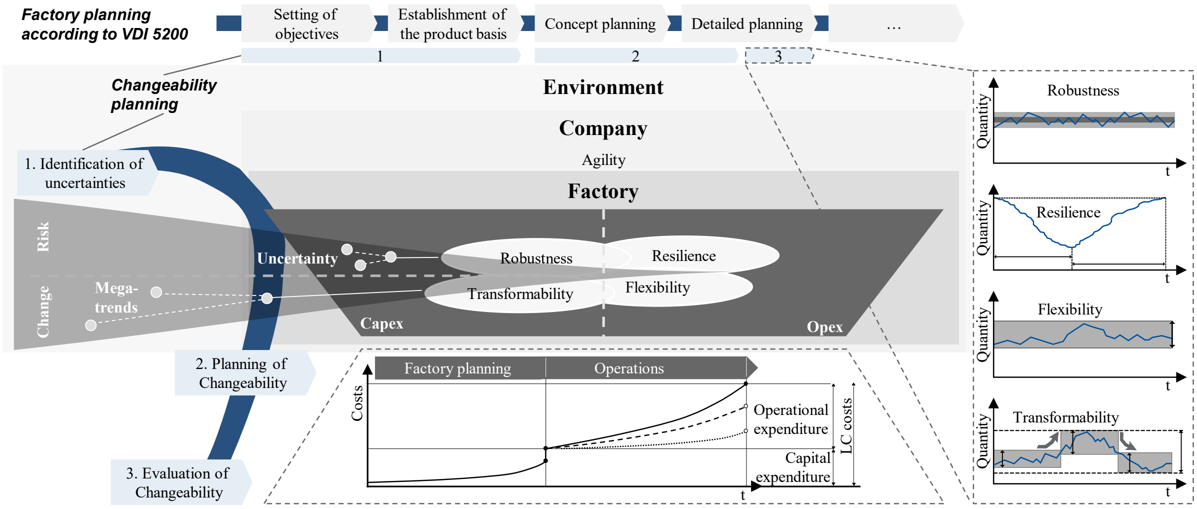


Figure 3: An overview of changeability planning in factory planning, based on [30,24,59]

The factory is part of a manufacturing company, which is located in an enterprise environment. In recent years, the degree of uncertainty has increased steadily both in the company's environment and in the factory system. On the one hand, this is expressed by long-term and continuous change in the form of superordinate megatrends, with effects on the company environment that are very difficult to predict. On the other hand, short-term disruptive events within the production system due to performance deviations inside or outside the factory are becoming more frequent. It is the task of factory planning to prepare the factory in the best possible way for the further course of the factory life cycle. For this purpose, different factory types or strategies of changeability are available to factory planning. Depending on the actual uncertainty present, the strategies of changeability capable of coping with the uncertainties most effectively over the further course of the factory life cycle must be designed. In order to be able to make such decisions in the context of factory planning, changeability planning must be integrated into factory planning. It is important that the uncertainties present in the use case are identified at the start of factory planning and considered from the very beginning. Subsequently, the changeability must be planned as part of the concept and detailed planning. An evaluation of the resulting factory planning variants is repeatedly carried out during this process, whereby changeability is usually only one of several factory objectives. In terms of changeability, it is important to evaluate whether the degree of changeability is sufficient for the predicted uncertainties in the course of the factory life cycle, so that the required performance level can be achieved, and which of the factory planning variants generates the lowest life cycle costs on top of that.

It can be concluded that resilience and robustness are rather mentioned in the literature in the context of short-term risks. Resilience ensures that when a disruption occurs, the performance curve only drops to a minimum that is tolerable for the system and then returns to the previous level in the shortest possible time. Robustness sets robustness limits to design a system to be able to operate under as many environmental conditions within the boundaries as possible. However, this is only optimal in a few cases, which is characterized by the narrow optimal range. Flexibility and transformability use static and dynamic flexibility corridors as classic characteristics of the capability for managing long-term change drivers. At the same time, implementing the dynamics of transformability and installing redundancies as part of robustness requires increased capital expenditures. The benefits of flexibility and resilience become more apparent during operations and consequently come along with increased operating expenditures.

4. Summary and Outlook

The increasingly complex factory systems are located in an extremely turbulent environment, resulting in numerous requirements and challenges. Therefore, the goal of this paper was to elaborate the planning process as well as different strategies of changeability. This knowledge will serve as a basis for identifying starting points for coping with diverse influences in the form of risks and change drivers. An analogy to the "two-factor theory according to Herzberg" accurately summarizes the findings: robustness and resilience are assigned to the area of dissatisfaction, meaning the negative effects on the factory. In contrast, transformability and agility refer to the area of satisfaction and enable the exploitation of opportunities and developments. Flexibility is ultimately found in both areas to some extent. Factory planning offers the opportunity to make decisions at an early stage, sometimes under considerable uncertainty, in order to prepare the factory for operation in the best possible way. However, mere knowledge of the factory types is not sufficient for this. They must be broken down to the inherent elements and processes in order to prepare them for risks and change drivers depending on their needs. The needs can be met in different ways. There will not only be one factory planning measure, resulting in different planning variants depending on the use case. Future research of the Institute of Production Systems and Logistics will work towards the quantitative evaluation of changeability over the factory life cycle. The impact of change drivers and risks over the factory lifecycle will be evaluated quantitatively in order to determine the right level of changeability in the context of economic efficiency. By knowing the behavior of the factory planning variants over the factory life cycle, they can be compared with each other. This should enable factory planners to select the factory planning variant that shows the lowest life cycle costs.

Acknowledgements

Partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 412409961. The authors wish to thank Antal Dér and Christoph Herrmann for the valuable discussions.

References

[1] Schenk, M., Wirth, S., Müller, E., 2010. Factory planning manual. Springer, Berlin, Heidelberg.

[2] Klemke, T., 2014. Planung der systemischen Wandlungsfähigkeit von Fabriken. Dissertation, Hannover.

[3] Stricker, N., 2016. Robustheit verketteter Produktionssysteme. Dissertation, Karlsruhe.

[4] Meyer, M., 2007. Logistisches Störungsmanagement in kundenverbrauchsorientierten Wertschöpfungsketten. Dissertation, Aachen.

[5] Hingst, L., Dér, A., Herrmann, C., Nyhuis, P., 2021. Evaluation of the influence of change drivers on the factory life cycle: (under revision). Procedia CIRP.

[6] Westkämper, E., Zahn, E. (Eds.), 2009. Wandlungsfähige Produktionsunternehmen. Springer-Verlag, Berlin, Heidelberg.

[7] Westkämper, E. (Ed.), 2013. Digitale Produktion. Springer Vieweg, Berlin, Heidelberg.

[8] Verein Deutscher Ingenieure, 2011. Factory planning - Planning procedures: VDI 5200.

[9] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2015. Handbook Factory Planning and Design. Springer, Berlin, Heidelberg.

[10] Herrmann, C., 2010. Ganzheitliches Life Cycle Management. Springer, Berlin, Heidelberg.

[11] Pawellek, G., 2014. Ganzheitliche Fabrikplanung, 2. Aufl. ed. Springer Vieweg, Berlin.

[12] Nielsen, L., Schmidt, C., Blume, S., Schmidt, M., Thiede, S., Nyhuis, P., Herrmann, C., 2016. Towards Quantitative Factory Life Cycle Evaluation. Procedia CIRP 55, 266–271.

[13] Schimmelpfeng, K., 2001. Risikomanagement in Industrieunternehmen, in: Götze, U., Henselmann, K., Mikus, B. (Eds.), Risikomanagement. Beiträge zur Unternehmensplanung. Springer-Verlag, Berlin, Heidelberg, pp. 277–297.

[14] Krebs, P., Müller, N., Reinhardt, S., Schellmann, H., Bredow, M. von, Reinhart, G., 2009. Ganzheitliche Risikobewertung für produzierende Unternehmen. ZWF 104, 174–181.

[15] Rogler, S., 2002. Risikomanagement im Industriebetrieb. Dissertation, Wiesbaden.

[16] Strohmeier, G., 2007. Ganzheitliches Risikomanagement in Industriebetrieben. Dissertation, Leoben.

[17] Wildemann, H., 2005. Management leistungswirtschaftlicher Risiken in der Produktion. ZWF 100, 187–191.

[18] Gottschalk, L., König, R., 2004. Bedarfsunsicherheiten durch flexible Kapazitäten beherrschen. ZWF 99, 715–719.

[19] Rühl, J.E., 2010. Monetäre Flexibilitäts- und Risikobewertung. Dissertation, Aachen.

[20] Wagner, C., 2012. Kontinuierliche Gestaltung skalierbarer Produktionsstufen. Dissertation, Hannover.

[21] Heinicke, M., 2017. Resilienzorientierte Beurteilung von Produktionsstrukturen. Dissertation, Magdeburg.

[22] Knüppel, K.M., 2016. Modellbasiertes Störungsmanagement in Produktionssystemen. Dissertation, Hannover.

[23] Ostgathe, M., Zäh, M.F., Grimmert, P., Schöler, T., 2011. Wissensbasiertes Störungsmanagement in Produktionsabläufen. ZWF 106, 838–843.

[24] Stricker, N., Lanza, G., 2014. The Concept of Robustness in Production Systems and its Correlation to Disturbances. Procedia CIRP 19, 87–92.

[25] Merz, M. Entwicklung einer indikatorenbasierten Methodik zur Vulnerabilitätsanalyse für die Bewertung von Risiken in der industriellen Produktion.

[26] Schmitt, R., Kostyszyn, K., 2015. Fehlerrisiken in der Produktion. wt Werkstatttechnik online 105 (11/12), 775–780.

[27] Pohl, J., 2013. Adaption von Produktionsstrukturen unter Berücksichtigung von Lebenszyklen. Dissertation, München.

[28] Unzeitig, W., 2014. Methodik zur frühen Fabrikplanung bei Unsicherheiten. Dissertation, Graz.

[29] Weig, S., 2008. Konzept eines integrierten Risikomanagements für die Ablauf- und Strukturgestaltung in Fabrikplanungsprojekten. Dissertation, München.

[30] Hawer, S., 2020. Planung veränderungsfähiger Fabrikstrukturen auf Basis unscharfer Daten. Dissertation, München.

[31] Heger, C.L., 2006. Bewertung der Wandlungsfähigkeit von Fabrikobjekten. Dissertation, Hannover.

[32] Löffler, C., 2011. Systematik der strategischen Strukturplanung für eine wandlungsfähige und vernetzte Produktion der variantenreichen Serienfertigung. Dissertation, Stuttgart.

[33] Möller, N., 2007. Bestimmung der Wirtschaftlichkeit wandlungsfähiger Produktionssysteme. Dissertation, München.

[34] Koch, J., Plehn, C., Reinhart, G., Zäh, M.F., 2014. Cycle Management for Continuous Manufacturing Planning, in: Zäh, M. (Ed.), Enabling manufacturing competitiveness and economic sustainability. Springer, Cham, pp. 9–12.

[35] Lübkemann, J., 2016. Ermittlung des Restrukturierungsbedarfs von Fabriken. Dissertation, Hannover.

[36] Benkamoun, N., 2017. Systemic design methodology for Changeable Manufacturing Systems. Dissertation, Aubière.

[37] Hernández, R., 2002. Systematik der Wandlungsfähigkeit in der Fabrikplanung. Dissertation, Hannover.

[38] Delbrügger, T., Döbbeler, F., Graefenstein, J., Lager, H., Lenz, L.T., Meißner, M., Müller, D., Regelmann, R., Scholz, D., Schumacher, C., Winkels, J., Wirtz, A., Zeidler, F., 2017. Anpassungsintelligenz von Fabriken im dynamischen und komplexen Umfeld 112, 364–368.

[39] Schenk, M., Seidel, H., Garrel, J. von, Grandt, C., 2012. Change Management in der Produktion – Flexibilisierungspotenziale nutzen., in: Rosenstiel, L., Hornstein, E. (Eds.), Change Management Praxisfälle. Springer, Berlin.

[40] Zäh, M.F., Müller, N. Prasch, M., Sudhoff, W., 2004. Methodik zur Erhöhung der Wandlungsfähigkeit von Produktionssystemen. ZWF 99, 173–177.

[41] Cucchiella, F., Gastaldi, M., 2006. Risk management in supply chain: a real option approach. Jnl of Manu Tech Mnagmnt 17 (6), 700–720.

[42] Wulf, S., 2011. Bewertung des Einflusses von Produkt- und Technologieveränderungen auf die Fabrik. Dissertation, Hannover.

[43] Sheffi, Y., Rice, J.B., 2005. A Supply Chain View of the Resilient Enterprise. MIT Sloan Management Review.

[44] Brünger, C., 2011. Nutzenkonsistente Risikopriorisierung: Die Risk-Map im Kontext rationaler Entscheidungen. Dissertation, 1. Aufl. ed. Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden, Paderborn.

[45] Westkämper, E., Löffler, C., 2016. Strategien der Produktion. Springer Vieweg, Berlin.

[46] Krause, D., Gebhardt, N., 2018. Methodische Entwicklung modularer Produktfamilien, 1. Auflage ed. Springer Vieweg, Berlin, Germany.

[47] Rosenkranz, F., Mißler-Behr, M., 2005. Unternehmensrisiken erkennen und managen. Springer, Berlin.

[48] Löhr, B., 2010. Integriertes Risikocontrolling für Industrieunternehmen. Dissertation, Frankfurt a.M.

[49] Deutsches Institut für Normung e.V., 2001. DIN6789-3: Dokumentensystematik: Änderungen von Dokumenten und Gegenständen allgemeine Anforderungen. Beuth Verlag,, Berlin.

[50] Becker, H., Langosch, I., 2016. Produktivität und Menschlichkeit, 5. Auflage ed. De Gruyter; De Gruyter Oldenbourg, Berlin.

[51] Fricke, E., Schulz, A.P., 2005. Design for changeability (DfC): Principles to enable changes in systems throughout their entire lifecycle. Syst. Engin. 8 (4).

[52] Schurig, M., Rabitsch, C., Ramsauer, C., 2014. Agile Produktion. ZWF 109, 956–959.

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

[54] Wiendahl, H.-P., ElMaraghy, 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.

[55] Tsourveloudis, N.C., Valavanis, K.P., 2002. On the Measurement of Enterprise Agility. Journal ofIntelligent and Robotic Systems 33, 329–342.

[56] Yusuf, Y.Y., Sarhadi, M., Gunasekaran, A., 1999. Agile manufacturing. International Journal of Production Economics 62 (1-2), 33–43.

[57] Rauch, E., Spena, P.R., Matt, D.T., 2019. Axiomatic design guidelines for the design of flexible and agile manufacturing and assembly systems for SMEs. International Journal on Interactive Design and Manufacturing 13 (1), 1–22.

[58] Vlajic, J.V., van der Vorst, J.G.A.J., Haijema, R., 2012. A framework for designing robust food supply chains. International Journal of Production Economics 137 (1), 176–189.

[59] Tierney, K., Bruneau, M., 2007. Conceptualizing and Measuring Resilience. TR news, 14–17.

[60] Trauth, D., 2019. Manufacturing Resilience : Towards Resilience Management in Production Engineering.

[61] Umsetzungsempfehlungen für das Zukunftsprojekt Industrie 4.0: Abschlussbericht des Arbeitskreises Industrie 4.0. Accessed 28 March 2021.

[62] Sethi, A.K., Sethi, S.P., 1990. Flexibility in manufacturing: A survey. The International Journal of Flexible Manufacturing Systems, 289–328.

Biography

Lennart Hingst, M.Sc. (*1989) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hanover since 2019. He previously studied technical logistics at the University of Duisburg-Essen (M.Sc.) and holds years of industry expertise.

Yeong-Bae Park, M.Sc. (*1994) has been a research associate in the specialist factory planning group at the Institute of Production Systems and Logistics (IFA) at the Leibniz University Hanover since 2019. He previously studied mechanical engineering at the TU Ilmenau (B.Sc.) and at the Karlsruhe Institute of Technology (M.Sc.).

Prof. Dr.-Ing. habil. Peter Nyhuis (*1957) has been head of the Institute of Production Systems and Logistics (IFA) at Leibniz University Hanover 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 (GAG3) since 2015. He is also a member of the German Academy of Technical Sciences acatech and the Scientific Society for Production Technology (WGP).

2nd Conference on Production Systems and Logistics

A Concept for Camera-based Classification of Load Carriers

Dimitrij-Marian Holm[1], Johannes Fottner[1]

[1]*Chair of Materials Handling, Material Flow, Logistics , Technical University of Munich, Garching, Germany*

Abstract

Due to growing environmental awareness, the Circular Economy and in particular the concept of Reverse Logistics (RL) are more and more becoming the focus of industry, yielding ecological as well as economic advantages. However, the successful implementation of the concepts requires that several challenges be met. One of the most common challenges is the lack of information within RL. One proposed solution is to use more Automatic Identification Systems (Auto-ID) to track returning goods and close the information gaps between RL participants. Currently available identification systems are often limited in their field of application, as they can be very expensive, require a huge change in current logistics processes or suffer from physical characteristics, such as electromagnetic absorption or limited visual contact. With this paper, we introduce our novel concept for load carrier classification and quantification using Time-of-Flight (ToF) cameras in combination with color images, beginning with a general overview of the system architecture and process structure. This is followed by an in-depth analysis of the process steps, starting with triggering camera records followed by image pre-processing, classification and finally a quantification of the loaded cargo.

Keywords

Circular Economy; Reverse Logistics; Classification; Load Carrier; Cargo Identification

1. Introduction

The discussion about the responsible use of natural resources has grown over the last decades. One major concept in addressing this issue is the Circular Economy [1]. It defines a production and consumption system that implements a circular life cycle for goods. This means not only that products, load carriers and packages are delivered from the manufacturer to the customer, but also that used goods are returned from the customer to the manufacturer for reuse or recycling, in a process called Reverse Logistics (RL) [2]. This process allows the sustainable usage of resources and components by returning them back to the commodity cycle [3].

Not only does the implementation of RL reduce the need for raw materials, companies can also gain economic benefits as well as social recognition [4]. Along with these benefits, RL faces several challenges, including social, political and economic barriers [5]. Other frequently mentioned major challenges are the lack of information, such as about the timing, quality or quantity of returned goods, and about the required technological systems needed to acquire this information [5]. If manufacturers implement Reverse Logistics, their production depends more and more on these kinds of information [6]. This paper contributes a technical concept for the classification and quantification of returned packages, especially load carriers, which can be easily adapted to current processes. The aim is to increase the amount of information which will improve knowledge about incoming and stored packages, enabling higher production certainty and better prediction of future material demands.

2. Related Work

2.1 Automatic Identification (Auto-ID)

In the Forward Supply Chain identification and tracking of goods are common tasks, mainly solved by using Auto-ID technologies. One of the most established and well used technologies are barcodes [7]. Even if they are easy to use and cheap, they have several disadvantages, such as the need for visual contact, the relatively short reading range [8] and susceptibility to environmental factors such as water and dirt [9]. On the other hand, Radio Frequency Identification (RFID) can solve the problem of visual contact and range, as it uses radio frequency for identification [10]. For the RL concept, RFID demonstrates good applicability for tracking returned products [11],[12]. While RFID seems to be a better choice than barcode-based identification systems, there are several drawbacks. RFID in general is much more expensive than a simple barcode. Besides tags, infrastructure for reading and processing the data is required [11], necessitating huge initial investments [13]. Further disadvantages result from physical characteristics. RFID uses radio frequency waves which are absorbed by liquids and reflected by metals, which reduces the utility of RFID in certain environments and industries [8],[14].

2.2 Machine Vision

With the breakthrough of modern Machine Learning (ML) methods such as Convolutional Neural Network (CNN), Machine Vision has experienced a huge acceleration in different areas such as Image Classification and Object Detection and Segmentation [15]. As ML methods are data driven, many datasets are available, either general [16],[17] or task-specific, such as autonomous driving [18] or remote sensing [19]. As logistics involves a large number of different objects and is associated with a high degree of continuous change [20], especially when customized load carriers are used, it is very difficult to find usable datasets. There are only a few datasets available, such as *Logistics Objects in Context (LOCO)* [21], which offers a wide range of general objects used in logistics, such as load carriers, pallets and forklifts, or, on the other hand, datasets that contain only one class (load carrier) [22].

3. Problem Description and Motivation

The lack of information within RL cannot be solved only by using currently available Auto-ID technologies, which leads to the need to consider alternative technologies. With the huge breakthroughs that modern machine learning techniques provide to machine vision, new approaches for identification are possible. However, huge datasets are needed for such approaches, and they are often not available within specific areas. For a better understanding, we emphasize an example using the German beverage industry. Here the end customer returns empty bottles as well as the load carriers, mainly crates, to the retailer, which delivers them to sorting companies and then back to the bottler. Future production therefore depends on the number of returned load carriers, which requires a process for classifying and counting. Most of the commonly used Auto-ID technologies such as RFID are not applicable in this industry because of the liquids and the high investment resulting from the huge number of crates in relation to the costs for installing tags. Auto-ID systems using optical identification, such as barcodes, are not applicable, because the composition of the transported goods is often changed by different processes like sorting. Old barcodes therefore need to be removed and new barcodes need to be attached to cargo at every process that changes composition, or else every load carrier would get a barcode, which is not possible due to the lack of visibility. Additionally, most beverage manufacturer design their own load carriers for product branding purposes and change the designs regularly, increasing the difficulty of data mirroring recently used load carriers. Due to the limitations of current Auto-ID systems, stocktaking is often performed by hand, resulting in an inaccurate and inefficient mapping of the inventory, and potentially affecting future production. With this paper, we contribute a concept to address both challenges, combining a method for fast data acquisition with the use of machine vision approaches to increase information density. This eliminates the need for manual stocktaking, reduces inventory differences and improves production forecasting. We aim to increase the scope of application and promote RL, mainly using cameras (ToF and RGB) and CNN for the classification of load carriers.

4. Concept Overview and Architecture

4.1 Use-Case

To explain our concept, we will focus on the load carrier transportation processes within a German beverage manufacturer. Beverage load carriers are regular and uniform, but vary in geometry depending on the type, so that carriers of one type can be easily stacked and transported on a pallet by forklifts. Depending on the forklift's size, it can simultaneously transport several pallets, which are called a bundle.

4.1 Planned Hardware Setup

For our setup we plan to use several sensors and computation units for different purposes. We will use two ultrasonic sensors with a measurement range of 4.5 meters to determine the distance between the forklift and the bundle and the height of the forks above the ground. For depth and color images, we will use a Microsoft Kinect V2, which combines color and a ToF camera in one single device. For process steering and data processing we will use an Nvidia Jetson NX Development Kit because of its small size compared to CPU performance and an integrated GPU to accelerate our CNN execution. All these components will be mounted on forklifts, so no external components are required.

4.2 Pipeline Architecture

For our concept, we designed a pipeline which can be subdivided into two parts. A general part provides a fast way to acquire and refine needed data (Triggering, Data Acquisition, Preprocessing) and two different modes – training (Labeling, Training) and operating (Classification, Quantification) – which can either be used for data labelling and training of neural networks or be used for classification and counting of load carriers (Figure 1).

Triggering and Data Acquisition: As the forklift approaches the load carriers, we measure the distance between the forklift and the load carriers. If the distance falls below a certain threshold, an event signal is sent to activate the camera and start data acquisition. As the forklift is driving towards the load carriers, the distance continues to decrease. If a second threshold is reached, data acquisition is stopped because the camera modules are too close to the load carriers and the field of view is restricted. The recorded data between start and stop are henceforth referred as time series. In our case our time series contain color and depth images, which are both needed for further processing. The end of data acquisition represents the start of the preprocessing task, which is realized by sending an event signal from data acquisition to preprocessing.

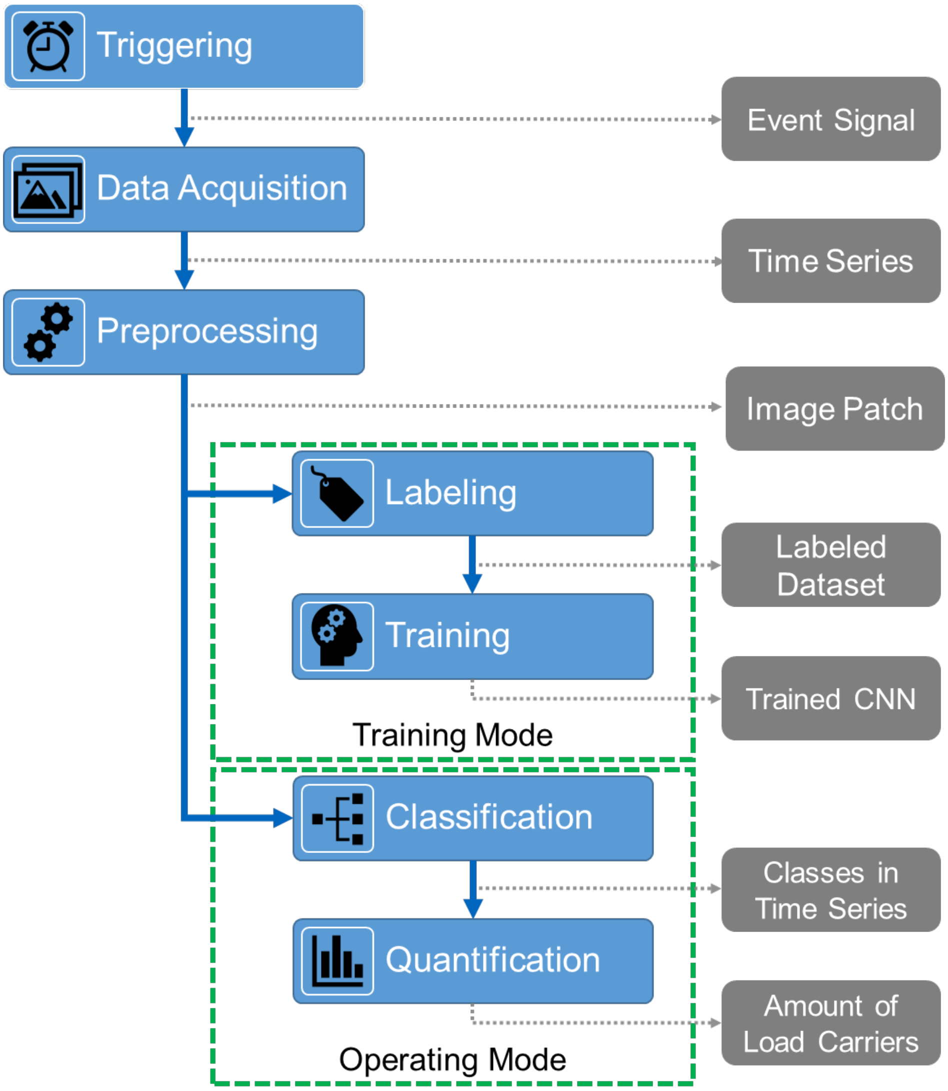


Figure 1: Pipeline overview.
Blue boxes describe the tasks while the blue arrows show transitions between the tasks.
Grey arrows and boxes refer to the results of the tasks.

Preprocessing: The preprocessing task will provide necessary data for both modes, training and operating. First, we need to specify the position of our transported load carriers and separate them from other load carriers to prevent misclassification. Therefore, the main task for preprocessing is to calculate the Region of Interest (ROI) in the color image in which the load carriers are contained (Figure 2).

Figure 2: Color image with marked RoI (red rectangle)

First, we identify specific key points within our depth image. As the load carriers are arranged as a cuboid, the most apparent key points are the four edges describing the front of the cuboid (Figure 3). Additional data about the cuboid are not used. Because our camera position is not restricted in height, the front is the only guaranteed visible part independent of the camera height.

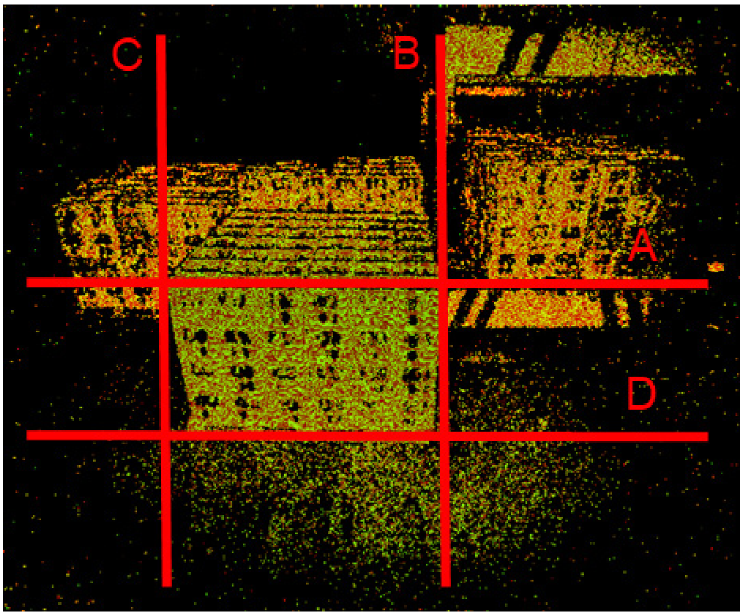


Figure 3: Depth image with marked edges.
A: upper edge, B: right edge, C: left edge, D: lower edge

By calculating the intersection of these edges, we find the corner points for our ROI, which are transformed from a depth image coordinate space to a color image coordinate space and used to crop the image (Figure 4). Since load carriers differ in size and we do not know the size of our selected load carriers, we subdivide our image into predefined areas which can handle different sizes to preserve geometrical and structural information. For example, if we choose one mask too small for our load carrier size, we cut off necessary information. Using a lager mask compensates for these information losses, but contains a small amount of additional information from neighbor load carriers. Therefore, we need to select our mask sizes precisely to find an equilibrium between the loss and additional information. In addition, this is a common practice to reduce problem of scale invariance from which CNN suffer [23]. We therefore cut our cropped image into smaller pieces by using different masks of predefined size (Figure 4). The resulting stack of small images is henceforth called a patch. These patches are used for both modes, training and operating. The patch is useful especially in the classification task to determine whether the bundle contains different types of load carriers. Depending on the current mode, patches are either stored for offline processes like labeling and training or used directly for classification and discarded afterwards.

Figure 4: Cropped color image and the resulting subdivided images using different mask sizes (red rectangle and blue rectangle)

Training mode: Training mode will be used for a fast generation of the necessary labeled data needed for the subsequent training of neural networks.

Labeling: Thanks to the preprocessing task, we can generate a huge number of cropped images automatically. In addition, since we are working with time series, we need to label only the time series instead of every image, since the images inherit the time series label.

Training: For training we will use Pytorch [24] as the framework for implementing and testing our CNNs for classification. Within Pytorch, several predefined CNNs such as ResNet [25] can be selected and trained with the generated patches. As our concept is developed to quickly switch to operating mode, object detection and segmentation are not considered because of the higher manual effort in labeling data.

Operating mode: With the operating mode, the picked load carriers can be classified and, by using additional information from the ultrasonic sensors, the amount of loaded cargo can be calculated.

Classification: Since we create our patch, we can now exploit the fact that every image contains only a small part of the original cropped image. Although we trained our CNN only with data created using bundles with single types, we can now classify every patch independently. By adding up the classification results of every single image, we can derive information about the whole bundle by evaluating the resulting distribution and making a statement whether there is only one or several types of load carriers within the bundle. In addition, if we use the position information of the patched images within our ROI and combine the information with our classification results, we can define areas within which certain load carriers occur.

Quantification: If the classification was successful and we use the lower edge and upper edge height, we can determine the total height of the cargo. Using the classification and area information, we are able to look up historical data related to our loaded cargo, e.g., the single load carrier height and the number of crates per layer, which finally can be used to estimate the total number of crates picked up by the forklift.

4.3 Restrictions

Due to the cargo loading process and the sensors mounted on forklifts, we can gain information from only one side of the bundle (Figure 2). Our pipeline can therefore estimate only the class and number of load carriers based on this data. Because this approach reduces the operating area, no further process adjustments or extra components need to be installed. In addition, we have further restrictions regarding the two modes of our concept.

<u>Training mode</u>: To accelerate labeling it is necessary that the loaded bundle contain only one class of load carriers. This allows us to label a whole time series with just one single label.

<u>Operating mode</u>: In contrast to the training mode restriction, different classes of load carriers can be handled within the bundle. However, with our restricted view of the bundle, we cannot detect mixed types outside the ROI. We therefore assume that, if our classification outcome predicts only one class, the whole bundle will be viewed as containing one type, otherwise the bundle is identified as mixed bundle. To quantify mixed bundles, an additional restriction comes into play. Because we have information about where different load carriers occur in our ROI, we can still estimate the number of load carriers if they are limited to pallets, e.g., every pallet contains only one type of load carrier. This allows us to look up historical data related to single-pallet cargo. If we cannot comply with this restriction, quantification is marked as impossible because of the high spatial mixing of different load carriers.

5. Conclusion and Future Work

Reverse Logistics as a part of the Circular Economy is a key concept for enabling a sustainable and environmentally friendly product lifecycle with additional benefits for companies willing to implement it. Beside the benefits, there are several challenges associated with Reverse Logistics, especially the lack of technical systems to acquire the information needed for successful operation. Current technologies such as Barcodes and RFID suffer from different physical, processual or economical limitations, e.g., electromagnetic absorption, limited visibility, major changes in current processes or high financial investments. In this paper, we presented a camera-based data processing pipeline for classification of load carriers to increase information density and overcome technical problems for manufacturers. Furthermore, our approach can be easily integrated into current processes without major changes, and the expense is scalable without huge initial investments, as no expensive external components are needed.

In future work we will implement this concept within a beverage facility to test the system under real conditions. We assume that the quality of images recorded by the camera system will be especially affected by environmental conditions such as rain, snow, the day-and-night cycle or fog, as the system will be used indoors and outdoors. Accordingly, we are defining our requirements and developing the software needed to realize our concept. Afterwards a test phase under real conditions is planned with an evaluation of the concept.

References

[1] Wautelet, T., 2018. The Concept of Circular Economy: its Origins and its Evolution.

[2] Brito, M.P. de, Dekker, R., 2004. A Framework for Reverse Logistics, in: Dekker, R., Fleischmann, M., Inderfurth, K., van Wassenhove, L.N. (Eds.), Reverse Logistics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–27.

[3] Korhonen, J., Honkasalo, A., Seppälä, J., 2018. Circular Economy: The Concept and its Limitations. Ecological Economics 143, 37–46.

[4] Thun, J.-H., Müller, A., 2009. An empirical analysis of green supply chain management in the German automotive industry. Bus. Strat. Env.

[5] Ravi, V., Shankar, R., 2005. Analysis of interactions among the barriers of reverse logistics. Technological Forecasting and Social Change 72 (8), 1011–1029.

[6] SRIVASTAVA, S., 2008. Network design for reverse logistics. Omega 36 (4), 535–548.

[7] McCathie, L., 2004. The advantages and disadvantages of barcodes and radio frequency identification in supply chain management.

[8] Helmus, M., 2009. RFID in der Baulogistik: Forschungsbericht zum Projekt "Integriertes Wertschöpfungsmodell mit RFID in der Bau- und Immobilienwirtschaft", 1. Aufl. ed. Vieweg+Teubner Verlag / GWV Fachverlage GmbH, Wiesbaden, Wiesbaden, 733 pp.

[9] Blumberg, D.F., 2004. Introduction to management of reverse logistics and closed loop supply chain processes. CRC, Boca Raton, Fla., London, 253 pp.

[10] Payaro, A., 2004. The Role of ICT in Reverse Logistics: A Hypothesis of RFID Implementation to Manage the Recovery Process, 7 pp.

[11] Plaza-Úbeda, J.A., Abad-Segura, E., Burgos-Jiménez, J. de, Boteva-Asenova, A., Belmonte-Ureña, L.J., 2021. Trends and New Challenges in the Green Supply Chain: The Reverse Logistics. Sustainability 13 (1), 331.

[12] Lee, C.K.M., Chan, T.M., 2009. Development of RFID-based Reverse Logistics System. Expert Systems with Applications 36 (5), 9299–9307.

[13] Vijayaraman, B.S., Osyk, B.A., 2006. An empirical study of RFID implementation in the warehousing industry. Int Jrnl Logistics Management 17 (1), 6–20.

[14] Koski, K., Koski, E., Virtanen, J., Björninen, T., Sydänheimo, L., Ukkonen, L., Elsherbeni, A.Z., 2012. Inkjet-printed passive UHF RFID tags: review and performance evaluation. Int J Adv Manuf Technol 62 (1-4), 167–182.

[15] Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., Alsaadi, F.E., 2017. A survey of deep neural network architectures and their applications. Neurocomputing 234, 11–26.

[16] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., 2009 - 2009. ImageNet: A large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops), Miami, FL. 20.06.2009 - 25.06.2009. IEEE, pp. 248–255.

[17] Everingham, M., van Gool, L., Williams, C.K.I., Winn, J., Zisserman, A., 2010. The Pascal Visual Object Classes (VOC) Challenge. Int J Comput Vis 88 (2), 303–338.

[18] Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research 32 (11), 1231–1237.

[19] Li, K., Wan, G., Cheng, G., Meng, L., Han, J., 2020. Object detection in optical remote sensing images: A survey and a new benchmark. ISPRS Journal of Photogrammetry and Remote Sensing 159, 296–307.

[20] Arnold, D., Isermann, H., Kuhn, A., Tempelmeier, H., Furmans, K., 2008. Handbuch Logistik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[21] Mayershofer, C., Holm, D.-M., Molter, B., Fottner, J. LOCO: Logistics Objects in Context. Proc. of 2020 IEEE International Conference on Machine Learning and Applications (ICMLA 2020).

[22] Poss, C., Irrenhauser, T., Prueglmeier, M., Goehring, D., Zoghlami, F., Salehi, V., 2019. Perceptionbased Intelligent Materialhandling in Industrial Logistics Environments, in: Proceedings of the 2019 11th International Conference on Computer and Automation Engineering - ICCAE 2019. the 2019 11th International Conference, Perth, WN, Australia. 23.02.2019 - 25.02.2019. ACM Press, New York, New York, USA, pp. 146–151.

[23] Gong, Y., Wang, L., Guo, R., Lazebnik, S., 2014. Multi-scale Orderless Pooling of Deep Convolutional Activation Features, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), Computer Vision – ECCV 2014, vol. 8695. Springer International Publishing, Cham, pp. 392–407.

[24] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library, 12 pp. http://arxiv.org/pdf/1912.01703v1.

[25] K. He, X. Zhang, S. Ren, J. Sun, 2016. Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.

Biography

Dimitrij-Marian Holm received his M.Sc. in robotics at the University of Applied Science Munich in 2018. Since 2018 he has been part of the research staff at the Chair of Materials Handling, Material Flow, Logistics, focusing on image processing and machine learning.

Johannes Fottner received a Dr.-Ing. degree in mechanical engineering from the Technical University of Munich (TUM), Munich, Germany in 2002. For the past 15 years, Johannes Fottner acted in different managing functions in the material-handling sector. Since 2016, he has been the head of the Chair of Materials Handling, Material Flow, Logistics with the TUM.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Situation Awareness Monitor And Liquidity Assessment For Enterprise Resilience Management

Julia-Anne Scholz[1], Nicole Oertwig[1], Caroline Delleske[2], Kai Fischer[2], Thomas Knothe[1], Alexander Stolz[2], Holger Kohl[1]

[1]*Fraunhofer Institute for Production Systems and Design Technology IPK, Berlin, Germany*
[2]*Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institut, EMI, Efringen-Kirchen, Germany*

Abstract

Currently, companies are becoming increasingly vulnerable towards unforeseen disruptive events. One example of such an event is the COVID-19 pandemic, in which many companies are facing existential difficulties. The sudden loss of suppliers and customers as well as frequent changes in regulations at short notice present companies with considerable challenges. Potential uncertainties and complexity complicate the planning and the implementation of measures. A fast and effective overview and the benefit of measures is usually missing. Situation awareness monitors have long provided orientation for first responders and rescue operators within the field of crisis management and can be just as relevant for companies.

Identifying, gathering, interpreting and visualizing external information from the corporate environment in relation to relevant disruptions remains an unsolved problem. In addition to the decision to implement effective measures before, during and after a crisis, an approach for the simulation of the disruptive impact and the effect of resilience measures on the company's performance as well as liquidity is required.

In this paper, a Situation Awareness Monitor (SAM) for resilience management is presented to address the aforementioned issues. First, the model-based IT architecture for the monitor is depicted. The enterprise model is coupled with an approach to assess critical processes of small and medium enterprises. Statistical models are applied to draw up the balance between single components, like input, output, supplier, storage and production. The result shows a continuous performance-based liquidity assessment to evaluate the resilience against disruptive events. Finally, single enhancement measures in alignment to the resilience phases – prepare, prevent, protect, respond and recover – can be analyzed.

Keywords

Business Resilience; Situation Awareness; Simulation; Crisis Management; Performance Monitoring

1. Introduction

Uncertain disruptive events and the changes of framework conditions are challenges especially for small and medium enterprises. Simultaneous disruptions to which a company is exposed or the potential adaptions before, during and after the crisis increase the complexity. This variety of information can lead to an overload of information management in companies. A fast provision of external and internal information before, during and after uncertain situations helps to derive effective and reliable solution strategies and decision-making aids. A Situation Awareness Monitor (SAM), taking significant resilience characteristics as well as liquidity simulation into account, enables an effective analysis and visualization.

The example of the COVID-19 pandemic shows that many companies have encountered existential difficulties due to the sudden loss of suppliers and customers as well as frequent short-term changes in regulations. The frequent increase of infections has exacerbated the uncertainty and made planning security and implementation of measures more difficult. By applying resilience management, the disrupting event can be conceived, analyzed and evaluated in advance at any moment.

The SAM for resilience management is able to combine incoming information and simulate the effects of disruptive events on a company's liquidity. This means that all changes can be displayed in a detailed overview of the current company situation, taking into account the time and the corresponding recalculation.

2. State of the art

2.1 Situation awareness

Since CLAUSEWITZ [1], the creation and use of SAMs have been of particular importance to emergency personnel of the military, police and health services for the success of networked action in crises. The starting and end point of the military command and control process as well as the police planning and decision-making process is always the situation picture, since it forms the foundation of the entire operation [2]. The relevant data must be identified, captured, structured and correlated. The second step is a structured analysis of the information obtained. The result leads to at least one, ideally to several decision options, within which the respective advantages and disadvantages are pointed out and weighed. On the basis of the decision taken, the implementation planning – which can be seen as a constantly recurring control loop – pursues the goal of a continuous actual-target comparison and makes it possible to react promptly to the changing situation. The goal of networked operations is to link the information, command and control systems from the infantryman's helmet camera to a military satellite as well as across all military organizations and units in order to aggregate as much relevant information as possible into a common situation picture. The information is then made available to the organizational units on a user-specific basis – for example, via an integrated display system in the infantryman's helmet. This enables individual decentralized systems and organizational units to work in a coordinated manner and within a shorter time. As flight control systems gradually became more automated and pilots started losing track of flight operations and aircraft status, concepts for situation awareness were further developed for the aviation industry [3]. One of the dominating definitions and theories was given by ENDSLEY: The three-level model approach describes the three steps for information processing, starting from "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and a projection of their status in the near future*" [4].

What manufacturing companies have in common with these sectors is the bias between a decision-maker's understandings of the system status – the company and its environment – versus the actual system status. This is mainly driven by the complexity of environmental influences, like natural disasters, climate change, geostrategic as well as economic policy influences, for example. In order to create a basis for further investigations and decisions for action, a large amount of information from a wide variety of sources must be collected and evaluated within a short time. This information must then be compared with existing findings, correlated with each other and communicated with responsible parties.

Hence, it is essential for companies to monitor the execution of their business in real time and to consistently have an overview of the current state of the company's liquidity and its facilities [5] [6]. Therefore, many companies today make use of so-called information systems, which are often summarized under the term management cockpit. However, the information processing is largely shaped by the hierarchical system of the automation pyramid [7]. This enables companies to provide operational data from process execution in production, to assign it to the corresponding order data and production specifications and to visualize it within the scope of performance analysis. Nevertheless, these instruments are just focusing on the impact of disruptions and are only partially suitable for the early planning of crisis containment measures and the assessment of their effects. They need extensions with regard to environmental influences as well as liquidity assessment.

2.2 Life cycle oriented resilience assessment

The reduction of potential risks is a central strategic challenge for small and medium enterprises. Classical crisis management calculates and estimates a possible disruption in advance, but does not make any specific adjustments in the further course of the disruption. SCHUMACHER shows that measures of a current crisis management are not suitable for small and medium enterprises [8]. Especially the COVID-19 pandemic showed that not many risk management systems were able to control risks within the supply chain or the maintenance of the business [9].

In contrast, a resilience analysis provides a holistic view of the situation. By continuously re-analyzing a disruption, the resilient approach provides, at any point in time (i.e. before, during and after a disruption), a reliable statement about the state and the effect of the disruption on the company as well as possible countermeasures and their positive effects on the company. With the help of resilience management, a specific assessment of the company's situation is carried out depending on the progress of the disruption.

While resilience describes the ability to react to a disturbance and to return to the desired initial state through suitable measures, the resilience cycle [10] deals explicitly with the evaluation of the disturbed system at any point in time. Therefore, the resilience cycle can be broken down into five interconnected phases: prepare, prevent, protect, respond and recover. The prepare-phase evaluates the system before the disruption occurs. Prevent, protect and respond are resilience phases that evaluate the system during the disruption, and recover initiates the time slot after the disruption including a new prepare-phase to continue the analysis into a cycle and preparing the system more robustly for a new occurrence by analyzing known events.

In addition, the Fraunhofer Resilience Evaluator (FReE) [11] was developed to support companies and systems to increase their resilience. For this purpose, a disruptive event is systematically analyzed and, with the subsequent evaluation, a company is enabled to successfully cope with a crisis [12]. The FReE tool gives a fast overview, but only a qualitative assessment concerning resilience.

2.3 Requirements and objectives for Situation Awareness Monitoring for resilience management

In order to create a SAM for resilience management and liquidity assessment for supporting tactical and strategic management decisions, it is of great importance to consider the previously mentioned open issues. In particular, the capability of monitoring the actual situation of the system in real time, collecting and evaluating the large amount of information from a wide variety of sources within a short time, correlating this information with existing findings and also communicating the results with responsible parties should be noted. Furthermore, the SAM for resilience management is characterized by its resilient, interactive, and interoperable performance. Specifically, resilience ensures situation awareness before, during and after a disruption. Interaction focuses on the communication between heterogeneous systems, and interoperability describes the clear visualization of relevant, reliable and up-to-date information, presented in a common situation picture. Thus, the functions to be supported by the system are:

1. Perform a continuous situation assessment and environment analysis to detect every change in status and obtain an early warning system. Furthermore, the simulation of the situation assessment monitor includes the impact of disruptions and of countermeasures, focusing on liquidity and performance.
2. Detect disruptions early and estimate potential effects through continuous analysis of trends, threats and deviations (e.g. delays in production).

3. Derive countermeasures and plan implementation support to obtain predictions of impacts of decisions and reliable predictions of environmental trends.
4. Implement the most appropriate countermeasures into the system.
5. Evaluate the changes and provide recovery.

These five steps correspond chronologically to the five phases of the resilience cycle (prepare, prevent, protect, respond, recover), ensuring a holistic resilient approach.

3. Situation Awareness Monitoring for resilience management

3.1 Target model

As explained above, the resilience cycle [10] serves the evaluation of the disturbed system throughout all five interconnected phases of a disruption. In Figure 1, the interdependencies between the activities and required resources in the resilience phases are described systematically with integrated enterprise modelling [13], supported by the modelling tool MO²GO [14].

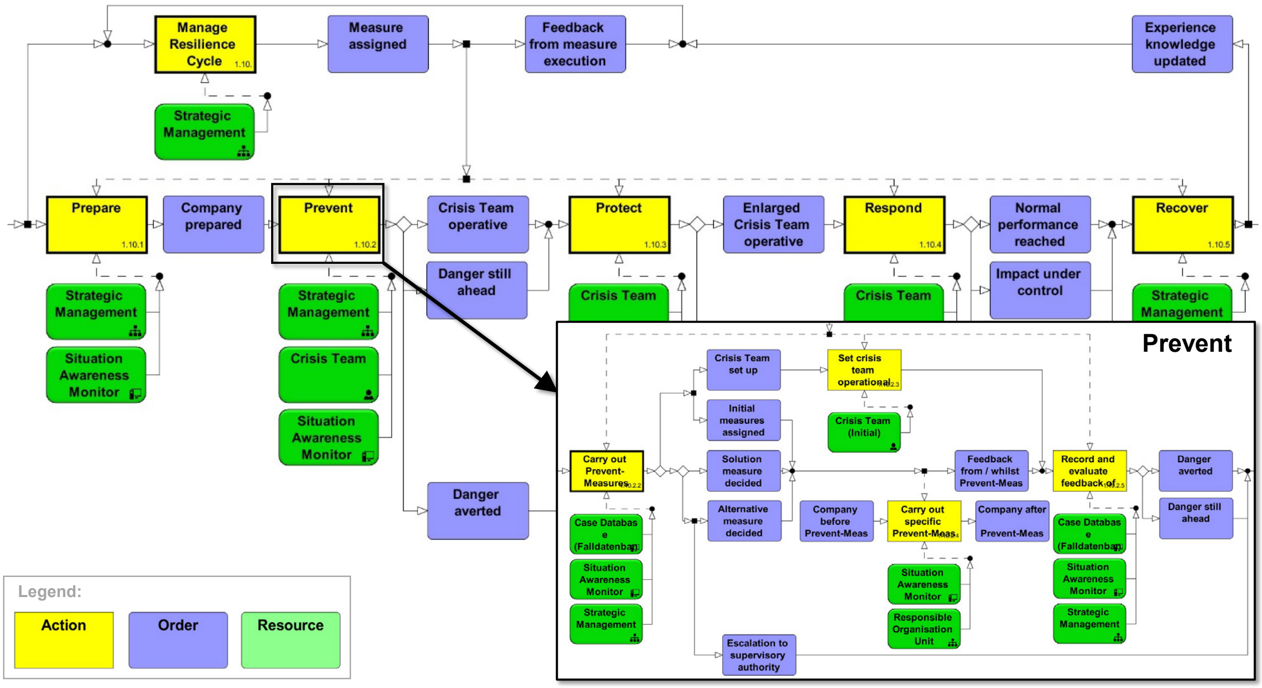


Figure 1: Process model excerpt from a reference model for holistic resilience management

The process "*Manage Resilience Cycle*" – which is depicted in the upper left corner of Figure 1 – triggers and controls the activities of the resilience cycle. It is an early warning system in the way that influences such as trends, external threats and internal deviations (e.g. delays in production) are continuously analyzed and their effect on critical performance parameters are evaluated based on simulated scenarios. If a potentially hazardous influence is identified, appropriate resilience measures are derived and assigned as controlling orders. In contrast to the other phases, activities in the prepare phase – such as ensuring the required redundancy by conducting employee trainings for multiple-qualification of employees – are carried out according to the continuous resilience assessment and are not only triggered by specific potentially hazardous influences. Based on the preparations of the company, prevent measures are carried out when a crisis becomes apparent and the previously defined crisis team is set operational. If the danger associated with the crisis could not be averted, protect measures must be carried out and the crisis team needs to be enlarged. In case the normal performance of the company could not be achieved, a rapid and well-organized response is necessary before the system enters the recover-phase in which the focus lies on stabilization and deriving lessons learned for improving the resilience cycle for future disruptions. Each phase includes a continuous evaluation of the measures being taken, which represents a feedback of the resilience cycle to the previously mentioned control process.

3.2 Model-based architecture

The reference model for holistic resilience management is an integrated part of the integrated enterprise model. As shown in Figure 2, the integrated enterprise model and the SAM for resilience management are the central components of the model-based architecture.

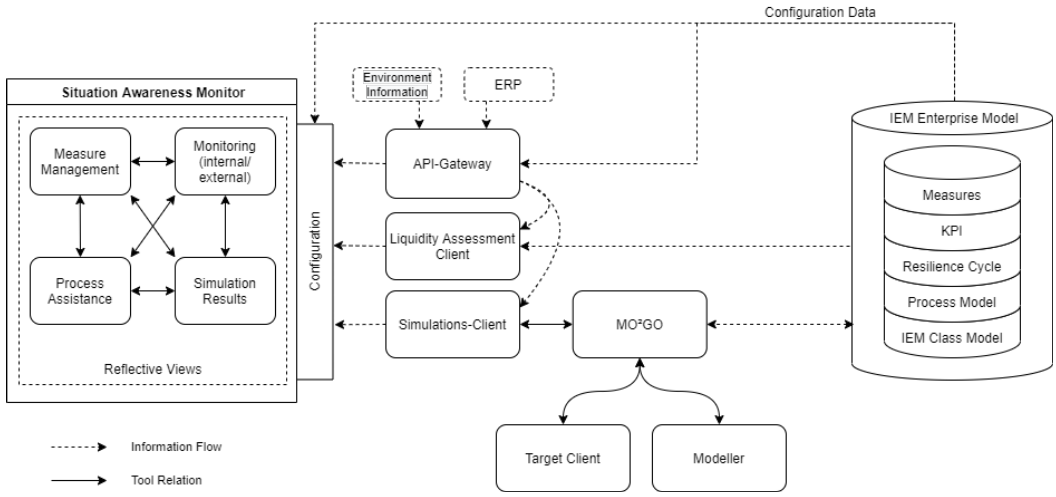


Figure 2: Model-based architecture for Situation Awareness Monitoring for resilience management

From the structural perspective, the integrated enterprise model combines the objects of an enterprise and its relationships with their behavioral aspects in processes [13]. Behavioral aspects can be, for instance, the definition of the key performance indicators (KPIs) in relationship to the enterprise objects and processes (e.g. duration, overall equipment effectiveness (OEE)). These model parts are used for the configuration of the SAM. In addition, the integrated enterprise model contains operational data as well (e.g. duration of a single process): KPI values and measures in all phases of the resilience cycle. The resilience measures are connected to the process definition in the model. Additional data from the corporate environment and internal IT systems (e.g. data from the enterprise resource planning system) are gathered by the application programming interface (API)-Gateway, which is configured by the integrated enterprise model as well. In the SAM, the information is displayed using four reflective view types: Measure Management, Monitoring of internal and external status of processes and resources (KPI-based), Process Assistance and Simulation Results. These view types can be instantiated by configuration. The API-Gateway is providing data for simulation and liquidity assessment. For the design of the integrated enterprise model, clients of the MO^{2}GO software [14] are used (e.g. modeler).

3.3 Resilience-based liquidity assessment

The model-based architecture helps to generate an abstracted overview of small and medium sized companies (SMEs) and to identify critical processes which give essential contributions to the liquidity of an SME. This section introduces an approach to assess the performance and the liquidity as basis for a resilience assessment. The left picture of Figure 3 shows a generic critical process of a manufacturing company including several components and their dependencies. The production depends on the number of orders, realized within the order management. The production further needs several components provided by the storage. An external dependence exists due to the component provision from appropriate suppliers. After the quality assurance, finished products contribute to the business volume, presented by the delivery process in the left picture of Figure 3.

For each component of the generalized model in the left picture of Figure 3, decisive parameters and their dependencies are developed and implemented into a tool for assessment. The result is a multi-dimensional model to describe the liquidity and the performance of a critical process. Single parts of the model are reserved with uncertainties. Examples are:

- supply risks concerning quality, quantity and delay,
- changing input based on a change of the market or the demand and
- changes within the production, based on the capacity or the performance.

In this approach, the theory of production logistics [15] is used to specify the resilience assessment with regard to the system to result in quantitative measures. The key figures of logistics are combined with probabilistic models to integrate potential risks, like failure rates or supply risks in combination with discrete event simulations, oriented to LANDTSHEER ET AL. [16] and CUBE ET AL. [17]. Insights of the reliability analysis of systems [18] are also considered. The combination of the different approaches allows a quantitative assessment of resilience. Performance and liquidity are derived as basic quantities.

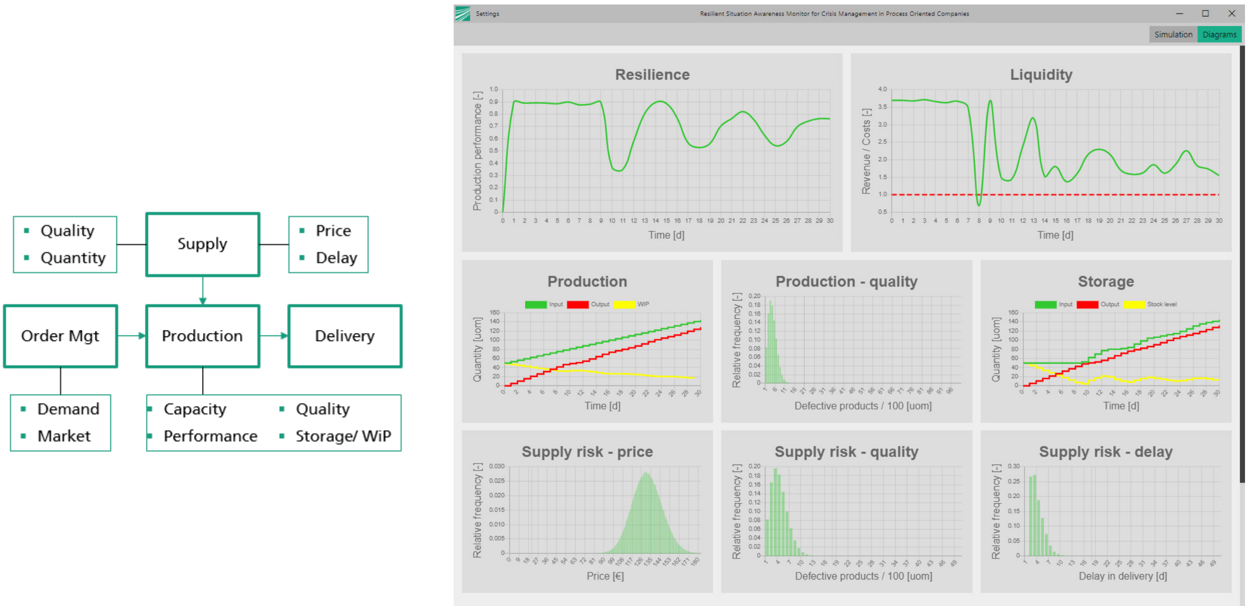


Figure 3: Left: Generic model to represent a critical process of an SME according to VDI 4499 [19]. Right: Overview of the graphical user interface to assess the resilience, the liquidity and further performance functions of a critical process.

The result of the multi-dimensional assessment is shown in the right picture of Figure 3. The user can simulate the resilience and the liquidity for an arbitrary number of shop calendar days. The impact and interdependence of certain parameters can be easily evaluated. The liquidity is expressed by the relation of revenue, based on finished goods, and the costs for employees, resources and the supplier. The performance target for the resilience quantification (upper left diagram) is based on the relation of maximum and current capacity of the production. The overall assessment scheme allows a quick overview concerning the consequences of disruptive events. Furthermore, the effectiveness of certain resilience phases, like preparation, protection, response and recovery can be evaluated.

4. Application in use cases

4.1 Application of the Situation Awareness Monitor for resilience management

The SAM for resilience management serves as an assistance tool for tactical and strategic decisions as well as the implementation of derived measures in the event of disruption and in daily operations. A prototype with a view generated for a company in the prepare phase, for example, is illustrated in Figure 4.

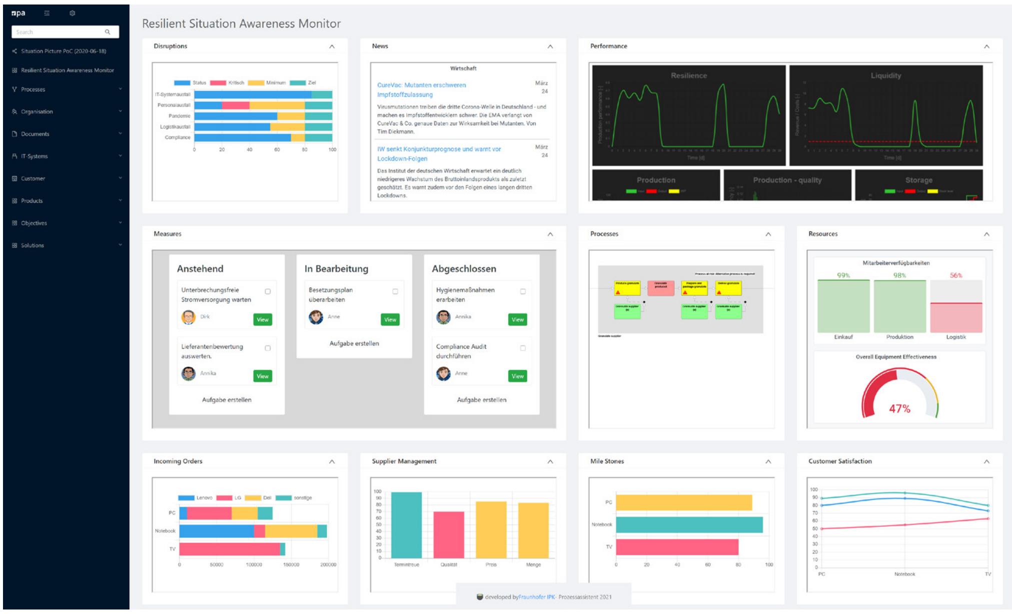


Figure 4: Prototype of the Situation Awareness Monitor for resilience management

Following the approach shown in Figure 1, recent news are continuously displayed in the upper left side of the prototype and form the basis for the corporate environment analysis, which represents the core element of the resilience cycle management process. In the widget left of the news ticker, the preparation status of the company concerning relevant disruptions is displayed. The associated resilience measures are managed in the Kanban board below. In the later phases of the resilience cycle, the Kanban board facilitates the implementation of measures – including setting the crisis team operational – and thus enables networked action. In preparation for a pandemic, the exemplary company illustrated in Figure 4 has developed a hygiene concept, which had previously been defined as the minimum preparation level that must be reached in order to be above the critical limit. In order to reach the target preparation level with regard to a pandemic, the staffing plan needs to be revised and updated. The corresponding resilience level as well as the liquidity level can be monitored and simulated in the multi-dimensional resilience-based liquidity assessment that is integrated into the upper right side of the prototype. Processes that are potentially affected by specific disruptions as well as the company's resources (e.g. staff availability and OEE) can be monitored on the center right side, which allows for a fast and targeted reallocation of resources in case of a disruption. In the lower section of the prototype, data of heterogeneous information systems (e.g. order management system, supply chain management system, manufacturing execution system, customer relation management system) is displayed for a comprehensive evaluation of the company's situation. This enables the management board to assess the feasibility of customer orders in case of a pandemic as well as to search and implement process alternatives.

4.2 SME liquidity during a disruptive event

The introduced assessment tool (as shown in Figure 3) generates a quick overview concerning the dynamic interaction of certain parameters of a critical process within an SME. Within the consideration of a disruptive event, the effects on the liquidity are analyzed. Taking the COVID-19 pandemic as an example, the effects are first seen in the suppliers and the customers. While many suppliers cannot provide reliable statements about the delivery status, the first impact on customers can be seen in the lockdown-induced retail closures. The action and information restrictions of these two elementary components of a critical process lead to massive limitations in the production of a manufacturing company. The affected parameters are primarily a drop in the economic indicator and order quantity, as well as an increase in storage time, storage space, and supplier pricing and delay calculation. Figure 5 shows the resulting liquidity in the left diagram, where a decreasing demand is considered. Furthermore, an increasing supply risk is realized by a larger uncertainty concerning the price and delay in delivery. After a few days, the liquidity decreases rapidly and losses are expected.

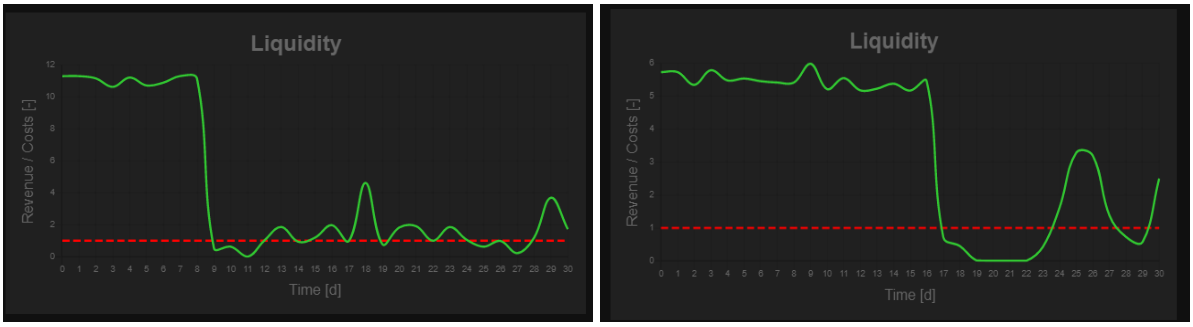


Figure 5: Left: time-dependent liquidity of an SME during a crisis situation (decreasing demand, increasing supply risk). Right: Similar disruption with consideration of short-time work as enhancement measure. For both simulations, random data is used for visualization.

Based on these impacts, suitable protective measures can be specifically derived. These are applied according to the category of the impact parameters in order to mitigate the disruption as far as possible at the point of origin. The right diagram of Figure 5 shows the result with the same disruption but the consideration of short-time work as countermeasure. The positive liquidity is reduced by a factor of two but extended on a longer time span. The decrease is induced to a delay in delivery by the supplier, but a fast recovery follows, based on the consideration of the enhancement measure.

5. Summary and Outlook

This paper provides an approach and a solution for an SAM for resilience management with an integrated resilience-based liquidity assessment. Due to its model-based architecture, the SAM and its reflective view types are generated by an integrated enterprise model, which enables a fast configuration of the SAM for various disruptions as well as daily operations and along the resilience phases. Furthermore, data from the corporate environment and internal IT systems is gathered and processed, which supports continuous performance and liquidity assessment. Moreover, the consequences of possible disruptive events as well as the effectiveness of resilience measures can be evaluated. Hence, the functions derived in Chapter 2.3 are supported by the SAM presented in this paper. The application example shows that the derived model considers the essential parameters to characterize the complex behavior of a small and medium enterprise. Further steps will focus on the validation of single elements, which is currently missing.

However, currently there are opportunities for improvement. So in the future, the views have to be adaptable directly in the SAM. When implementing the appropriate countermeasures, for example, adjustments need to be made in the integrated enterprise model, which then generates the SAM views. Therefore, further research is needed on the technical level in order to agilely enable the user to make adjustments in the SAM, which are reflected in the integrated enterprise model and vice versa. Concerning the methodological approach, further research should be conducted on improving the accuracy of the SAM in order to assess the completeness and the consistency plausibility. Likewise, it is to be evaluated under which circumstances the parameter-based liquidity assessment is preferred to the event-discrete simulation.

Another aspect to be taken into account is the relationship between the introduced model-based resilience monitoring and the performance management using electronic performance boards on the production shop floor. On the one hand, a number of artefacts are addressed on both systems, e.g. output rate, sickness level, measure definition and monitoring. On the other hand, some elements are currently suitable in one of the systems only (e.g. external information, events in the SAM). In order to avoid complexity and redundant systems, further research has to be performed for synchronization and contextualization. So for instance, advanced performance boards can be enabled to integrate necessary views and functions as required along the resilience phases (e.g. communication channels to crisis management board).

References

[1] Clausewitz, C. von, 2019. Clausewitz - Vom Kriege: Militärische Führung und Strategie, 1st ed. epubli, Berlin.

[2] Schaub, H., 2016. Vernetzte Operationsführung zur Unterstützung militärischer Stäbe, in: Hofinger, G., Heimann, R. (Eds.), Handbuch Stabsarbeit. Führungs- und Krisenstäbe in Einsatzorganisationen, Behörden und Unternehmen. Springer, Berlin, Heidelberg, pp. 291–295.

[3] Stanton, N., Chambers, P., Piggott, J., 2001. Situational awareness and safety. Safety Science 39 (3), 189–204.

[4] Endsley, M.R., 1995. Toward a Theory of Situation Awareness in Dynamic Systems. Human Factors 37 (1), 32–64.

[5] Kung, P., Hagen, C., Rodel, M., Seifert, S., 2005. Business Process Monitoring & Measurement in a Large Bank: Challenges and Selected Approaches, in: Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA'05), Copenhagen. IEEE, pp. 955–961.

[6] Harmon, P., Wolf, C., 2014. The State of Business Process Management 2014. Business Process Trends. http://www.bptrends.com/bpt/wp-content/uploads/BPTrends-State-of-BPM-Survey-Report.pdf. Accessed 10 March 2021.

[7] Heinrich, B., Linke, P., Glöckler, M., 2017. Grundlagen Automatisierung: Sensorik, Regelung, Steuerung, 2nd ed. Springer Vieweg, Wiesbaden.

[8] Schumacher, J., 2020. Coronavirus-Krise: Die langfristigen Folgen für KMU. https://unternehmer.de/management-people-skills/250243-coronavirus-krise-folgen-fur-kmu. Accessed 31 March 2021.

[9] KPMG, 2020. Unternehmensüberwachung – Welche Entwicklungen jetzt im Fokus stehen. CGP Das Governance Magazin 9.

[10] Thoma, K. (Ed.), 2014. Resilien-Tech: "Resilience-by-Design": Strategie für die technologischen Zukunftsthemen. Utz, München.

[11] Hiller, D., Finger, J., 2021. Fraunhofer Resilience Evaluator. Fraunhofer Ernst-Mach-Institut. https://www.resilienz.leute.server.de/free/. Accessed 10 March 2021.

[12] Fraunhofer Ernst-Mach-Institut, 2021. Resilienz als Schlüssel, um Krisen erfolgreich zu bewältigen. Fraunhofer Ernst-Mach-Institut. https://www.emi.fraunhofer.de/de/geschaeftsfelder/sicherheit/forschung/Fraunhofer-Resilience-Evaluator-FReE.html. Accessed 10 March 2021.

[13] Spur, G., Mertins, K., Jochem, R., 1993. Integrierte Unternehmensmodellierung, 1st ed. Beuth, Berlin.

[14] Fraunhofer Institute for Production Systems and Design Technology IPK. MO²GO: Enterprise Modelling, Analysis and Optimisation. https://moogo.ipk.fraunhofer.de/. Accessed 16 March 2021.

[15] Nyhuis, P., Wiendahl, H.-P., 2009. Fundamentals of production logistics: Theory, tools and applications. Springer, Berlin, Heidelberg, XIX, 312 S.

[16] Landtsheer, R. de, Ospina, G., Massonet, P., Ponsard, C., Printz, S., Härtel, L., Cube, J.P. von, 2016. A Discrete Event Simulation Approach for Quantifying Risks in Manufacturing Processes, in: ICORES. Proceedings of 5th the International Conference on Operations Research and Enterprise Systems, Rome. SCITEPRESS - Science and Technolohy Publications Lda, Setúbal, pp. 313–322.

[17] Cube, P. von, Härtel, L., Schmitt, R., Ponsard, C., Massonet, P., Landtsheer, R. de, Ospina, G., Printz, S., Jeschke, S., 2016. Monetary Quantification of Supply Risks of Manufacturing Enterprises - Discrete Event Simulation Based Approach, in: Editorial 49th CIRP International Conference on Manufacturing Systems (CIRP CMS), Stuttgart. Elsevier, Amsterdam, pp. 164–170.

[18] Straub, D., 2018. Lecture Notes in Engineering Risk Analysis. TU München, München.

[19] Verein Deutscher Ingenieure e. V., 2008. VDI 4499 Part 1: Digital factory. Accessed 31 March 2021.

Biography

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.

Nicole Oertwig (*1984) is Deputy Head of the Business Process And Factory Management Department at the Fraunhofer Institute for Production Systems and Design Technology (IPK). Her research focuses on visualization solutions for Industrie 4.0 and resilient business alignment. She has studied business Mathematics with a specialization in risk management at the HTW Berlin.

Caroline Delleske (*1991) is a research associate at the Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institut (EMI) since 2019. She is a researcher for robustness and resilience analyses. Previously, she has studied Physics at the Goethe University in Frankfurt.

Dr.-Ing. Kai Fischer (*1982) is the head of the group Robustness And Resilience Analysis at the Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institut (EMI) in Freiburg. He received a diploma in Applied Mathematics, a Master degree in Building Physics and a PhD in Engineering Sciences.

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.

Prof. Dr.-Ing. Alexander Stolz (*1975) is the chair for Resilience Engineering for Technical Systems at the Faculty of Engineering, Department for Sustainable Systems Engineering (INATECH) University of Freiburg and Head of the Department for Safety Technology and Protective Structures at the Fraunhofer Institute for High-Speed Dynamics, Ernst-Mach-Institut (EMI).

Prof. Dr.-Ing. Holger Kohl (*1971) is the deputy director of the Fraunhofer Institute for Production Systems and Design Technology (IPK) and a professor at the Technical University of Berlin, holding the chair for corporate sustainable development. He is chair of the annual Global Conference on Sustainable Manufacturing.

2nd Conference on Production Systems and Logistics

The Barriers for Iranian Retailers to Outsource Their 3PL Activities

Daniel Sahebi, Petri Helo

Production department, University of Vaasa, Finland

Abstract

The importance of supply chain management and its related complex activities has been widely noted recently. Many companies desire to enter the international market in today's competitive environment, and logistics could be a significant strategic tool as well as information source. On the one hand, the emergence of third-party logistic companies have solved many of their entry problems and accordingly gained a great deal of attention. On the other hand, there are many barriers that, despite great need, stop companies from trusting and outsourcing their responsibilities to another party. This paper has surveyed retailers about their problems with outsourcing their logistics activities. While some factors such as a problem in the coordination are not considered important by the retailers, there are some interesting issues like law barriers and the absence of good review that would stop them choosing a third party to handle logistic activities.

Keywords

Supply chain management; Third party logistics; Outsourcing; Production

1. Introduction

Supply chain process includes a multitude of actions from creation to consumption. In the supply chain, the flow of goods is from the production chain to the consumption chain while the flow of money is from the consumption chain to the production. The flow of information and value, however, is mutual. The competitiveness of a product depends on the performance of these three flows in the supply chain organization. The competition rate is growing quite rapidly in the Iranian market, and it has shifted from the level of individual firms to the supply chain and prevailing business environment throughout the industry. There are various factors as to why the concept of Third-party Logistics (3PL) has not been fully taken advantage of in Iran. In this research, we firstly look at the concept of 3PL as a whole. Then we consider the challenges facing third-party companies in Iran. Finally, this paper explores the main barriers for retailers, in Tehran, the capital city of Iran, to outsource their supply chain activities to a third party and to choose a reliable 3PL partner.

2. Literature review

2.1 Types of logistics service providers

Several definitions of the supply chain have been provided. Generally, the supply chain is the set of all firms that directly or indirectly contribute to accomplishing the consumer's demands [2]. In recent decades, manufacturers have found that, to gain competitive advantage and to survive in the global competition, they should focus their attention on production, costs reduction, improving the service level, and outsourcing their logistics activities. The establishment of companies providing logistics services and assigning logistics operations to them is one of the most important measures that can help industry and enterprise owners [10]. During the previous decades, there have been a vast number of definitions with regards to the nature of 3PL as a concept and the providers of these services. Some of these definitions are mentioned below.

According to Lieb [23], "3PL is concerned with the use of external companies to carry out logistic tasks which have been previously carried out within the firm. The tasks carried out by the third party can cover the whole logistics function or specific tasks within the process". Berglund et al. [4] (1999, p.59) defines Third party logistics as following:

"functions taken care of by an external service provider on behalf of another firm which is consisted of at least the management and execution of transportation and warehousing processes."

Furthermore, the duration of cooperation between the two firms should surpass one full year for the service provider to be known as a 3PL company, otherwise it is considered as 'arm's length' sourcing of transportation or warehousing. Evangelista and Sweeney [12] (2006, p. 56) define the outsourcing in logistics as follows:

"Third party logistics are functions which are taken care of by a logistic service provider on behalf of a shipper and is consisted of, at least, the process of transportation." The concept of service offering in 3PL can also include various functions such as warehousing and the management of inventory, in addition to activities linked to value added supply chain and information. The 3PL industry is quite vast in size, therefore it is worth studying on its own. The academic attention in the field of 3PL has been increasing in the past decade. This is basically due to the increasing interest of firms in outsourcing most of their functions which are non-core activities [33].

Various organisations, with the outsourcing of many of their public services, reduce the need for investment and asset creation within their firms [38]. As an example, an organization with outsourcing of their warehousing to 3PLs does not need to invest heavily in warehouses and refrigeration facilities. Meanwhile, the organization can use a 3PL service to direct a part of its capital rather than deploying assets to advantaged sectors. As an example, the organization, with the assistance of the transport fleet, no longer needs to invest in the purchase of transportation vehicles. 3PLs assists the company to maximize their capital and properties for storage. Failing to do so would lead to the loss of many corporate savings, because the warehouse or refrigerator of the firm is not always fully utilized. Therefore, part of these assets will be underutilised and will even result in fixed and current costs (repair, maintenance, old age, depreciation, etc.). However, with the outsourcing of services to 3PLs, the costs mentioned above are withdrawn from the responsibilities of the organization [17].

Deciding upon the most eligible 3PL service provider has become a rather demanding task for many international companies. The selection of the most suitable service provider begins with correct establishment of the right criteria for the identification and evaluation of the right 3PL service provider company. These firms should have the highest potential for meeting the requirements and standards of service buyers [5]. Complex or multi criteria decisions related to the selection of the service providers with the most potential, is mainly related to the reasoning and the experience of the management team in charge of the selection process [9], [19]. Taking into consideration various multifaceted decision-making issues, hand picking of 3PLs service providers can be a multiple-criteria decision-making task (MCDM) [26]. While evaluating 3PL service providers, managers usually face a certain level of uncertainty, and the amount of input data is not accurately clear [28].

As mentioned earlier, the decision-making process is always accompanied with uncertainty. The uncertainty factors are usually expressed in the form of interval numbers, rough numbers, fuzzy numbers, and grey numbers [17]. These are mostly suited for the presentation of uncertainty, linked to qualitative criteria by taking advantage of the linguistic scales and determining how reliable the expert`s evaluation is [35]. The idea of applying algorithms in the process of decision-making supported by the interval approach, mainly implies that interval numbers are being used for the presentation of attribute values [13]. However, it is a rather demanding task to pinpoint the limitation of interval numbers, due to the fact that they are mainly based upon the intuition, experience and the perception of the managers in charge of the decision-making processes [29].

Studies show that the service level of 3PL companies is increasing every year. Also, 3PL users utilize participatory collaboration which serves the third-party logistics service companies to improve performance and enhance their level of collaborations [14].

Based on the concept of dynamic capability, to gain a long-lasting competitive advantage, companies which are competing in active markets, should be obtaining, intergrading, reconfiguring, and releasing resources [18], [24]. These factors in dynamic capabilities, precursor various norms and routines which companies develop for the purpose of value creation, aiming at better and more efficient responses to various forces within the market. The learning orientation of a company can directly influence their efficiency in adapting both effective and timely routines [36]. Although researchers have conducted various studies on learning orientation and its role on competitive advantage, [8], [14], [16], [32], [39], its influence on outsourcing of logistics remains largely unexplored [7]. Panayides [30] provided a substructure on how a company`s learning could lead to a successful link between logistics service providers and their service buyers. However, for 3PL companies, knowledge sharing, as a leading process for organization learning, is one of the biggest barriers for adoption of such services [31].

Due to reduced product costs, efficient logistics infrastructure leads to the prosperity of both producer and consumer. On the other hand, improving logistics in geographically suitable countries help them earn more income and increase employment rates. Currently, the improvement of the status of countries in the field of logistics has become one of the major goals of economic development. Since the inappropriate situation of the logistics network in each country increases business costs and reduces the ability of countries, especially developing countries, to join the global supply chain, policymakers consider essentially the implementation of coherent and sustainable policies to strengthen and integrate the supply chain and the improvement of the logistics situation not only in the best countries in the field of logistics, but also for emerging countries as an engine for growth [25].

2.2 Current situation in the 3PLs services market in the world

Globalization as a concept has fundamentally and permanently altered the business environment. In response to this recent economic norm, various producers and suppliers have recognised 3PL companies as crucial players in the process of the supply chain. As dedicated logistic specialists, these companies enable their customers to concentrate on their core functions and areas of expertise [34], [40] and this focus ultimately lead to the improvement of the competitive advantage of the firm [1], [6], [20].

The concept of customer and supplier relations has been challenged dramatically by the interdependence created due to outsourcing [22], [37]. Based on the seminal work of Morgan and Hunt [27], competition is included as the main characteristic of any free market. However, one can claim that commitment and trust are absolute necessities for the maintenance of a successful committed business relationship. In these convoluted and reciprocal relationships, the concept of success or failure depends on the presence of honest collaboration and the lack of opportunistic behaviours. Based on various literature, trust is one of the main factors in the promotion of cooperation and leads to a higher rate of data sharing amongst various interdependent parties [11], [15].

2.3 Challenges facing third-party companies in Iran

In previous sections, the position of 3PL companies in improving business conditions has been reviewed. It highlights the importance of developing these companies in Iran. However, the specific conditions and characteristics of Iran, examples of which are given below, make the issue more relevant:

- Positioning in a strategic place due to geographical location: These conditions have enabled Iran to become a logistical link between Asia and Europe, on the one hand, and a communication link between the countries of Central Asia and the Free Seas on the other. This special position is an advantage, which if properly addressed and effectively used, could generate revenue way beyond the revenue produced from oil [11].
- Definite need for active participation in international markets: Due to the policy of non-reliance on oil revenues, which is one of the main strategies for the country's economic development, it is necessary that domestic products are offered at world-class quality and price. This is rather an impossible task to accomplish without proper and optimal logistics infrastructure [21].
- The need to reduce prices: Due to recent economic trends in the country and rising prices of energy carriers, and consequently an increase in raw material prices, the cost of products and services has increased significantly. Undoubtedly, by optimizing energy consumption, raw materials or, in other words, increasing resource efficiency, the impact of these factors could be significantly reduced. This issue is the main priority in various industries and organizations [2].
- Supply and supplier related issue, fragmentation, reliability in delivery, flow of information, systems for the management of quality and inventory, disruption of customer service, processes of research and development, workers´ skill level, planning, costs related to production and waste management, fiscal management, currency rates, logistic, demand, and regulations are some of the main topics which are considered the many risks for manufacturing companies in Iran [21].

2.3.1 Sanctions against Iran

Sanctions against are imposed by global powers to punish Iranian government or to persuade this country to give up some activities. Most of the sanctions against Iran are formed after the occupation of the U.S embassy in Tehran in 1979 and they are intensified with the challenge of Iran's nuclear program. Iran's atomic program has faced many sanctions in 2016, after the referral of the case from the atomic agency to the security council. Sanctions against Iran can be divided into two-four general groups based on the countries that impose the sanctions: multilateral sanctions such as sanctions of the security council, sanctions of the European union, unilateral sanctions of different countries such as U.S and sanctions of the U.S congress. In the current age, the economy is considered as an important basis of wealth and power production and a basis to reach growth and economic development. This has attracted the attention of Iranian policymaker and there have been many economic plans to reach economic development. In this regard, economic sanctions are the main barriers to reach this goal. Sanctions have targeted a large section of Iran's economy and have limited access to financial resources and currency, have reduced investment, have increased inflation and unemployment, and finally have reduced economic growth.

Iran's economy will face a burdensome crisis under the influence of the U.S sanctions if the European Union does not cooperate with Iran (and it seems E.U is not willing to cooperate). Iran's oil export will face negative growth. The policy of oil for goods is the government's policy with other nations. Of course, given the government's policy to get dollar prices closer it's a real value, we will witness higher export for full national production that can somehow cover negative growth of the oil productions. Unfortunately, Iran will again face two-digit inflation. The concern here is that the country will again enter stagnation and the only possible solution is to use policies to increase export [31].

2.3.2 Logistics in Iran

There are 3 major players when it comes to logistics in Iran. These players are Government and governmental related organizations which hold 60% of the shares. Academia holding 25% and operating companies which hold 15% of the shares. Each of the mentioned sections contain several subsections which are stated below:

Government

- Road Development Ministry
- Organization of transportation and maintenance of roads
- maritime corporation
- Civil aviation corporation
- Islamic Republic of Iran´s railway

Operations

Shipping Lines of the Islamic Republic of Iran

- Port of Sina
- Middle East´s Tidewater
- TukaRail Co
- International Transportation Co. of Persian Gulf
- Iran`s Tosse Tarabari
- International Transportation of Iranian Chain
- Kaveh Co
- Sepahan Co
- Iran Air corporation
- Logistics Holding of Khadem
- Aras Asia Seir Co

Academia

- Logistics and Supply Chain Society of Iran
- Imam Hossein University Logistics Study and Research Centre
- Trade Studies and Research institute

Unfortunately, due to the political structure of the country, even though the government of Iran holds only 60% of the shares of the logistic industry, it still has a very high influence on the rest of the shareholders. Iran has many advantages and potential in the field of logistics, of which the most important are the Middle East and the main five major corridors of transit and international transportation, the use of all modes of transportation, access to high seas, having blue shores and extensive land borders, having numerous ports and extensive road network. Due to these advantages, the country has the ability to become one of the major commerce centres in the region and even the world, but unfortunately, so far, the development of logistics in the country is not been significantly taken into consideration. Looking at Iran's position in the indicators related to the field of logistics is a good express of the country's weakness in this regard and the lack of development of the logistics infrastructure.

2.3.3 International logistics in Iran

The Logistic Performance Index (LPI) is focused on a global survey of operators (global freight forwarders and express carriers) who provide input on the "logistic friendliness" of the countries in which they work and trade. They incorporate in-depth knowledge of the countries in which they work, comparative evaluations of other countries in which they deal, and global logistics expertise. Quantitative data on the success of main components in the logistics chain in the country of work is combined with feedback from operators. As a result, the LPI includes both qualitative and quantitative indicators, and it aids in the creation of logistic friendliness profiles for these countries. It assesses success across a country's logistics supply chain and provides two perspectives: global and domestic. The LPI overall score reflects expectations of a country's logistics based on the reliability of customs clearance, the consistency of trade and transportation-related facilities, the convenience of organizing competitively priced imports, the quality of logistics providers, the ability to track and trace consignments, and the pace at which shipments arrive at the consignee on time. A higher score indicates better results. The index varies from 1 to 5, with a higher score indicating better performance [25].

Table 1. Logistic Performance Index in Iran and other sectors of world

COUNTRY	LPI SCORE	CUSTOMS	INFRASTRUCTURE	INTERNATIONAL SHIPMENTS	LOGISTICS COMPETENCE	TRACKING & TRACING	TIMELINESS
Region: Europe & Central Asia	3.03	3.16	3.14	3.17	3.24	3.62	3.23
Region: East Asia & Pacific	2.98	3.02	3.08	3.07	3.12	3.54	3.14
Region: Middle East & North Africa	2.6	2.78	2.96	2.81	2.86	3.29	2.89
Region: Latin America & Caribbean	2.48	2.46	2.69	2.6	2.67	3.05	2.66
Region: South Asia	2.42	2.45	2.68	2.56	2.56	3.03	2.62
Region: Sub-Saharan Africa	2.36	2.29	2.49	2.42	2.39	2.84	2.47
Iran, Islamic Rep.	2.6	2.33	2.67	2.67	2.67	2.44	2.81

The followings are the six main dimensions which are used to evaluate countries:

1. Border control agencies, like customs, are efficient in their clearance processes (i.e., speed, simplicity, and predictability of formalities).
2. Infrastructure for commerce and transportation (e.g., ports, railroads, roads, and information technology).
3. Ease of coordinating low-cost shipments.

4. Logistics facilities (e.g., transport providers, customs brokers) competence and quality.
5. Consignment tracking and tracing capabilities.
6. Shipments arriving at their destination on time or ahead of schedule.

Iran is ranked 112th in World Bank survey that examines logistics conditions and infrastructures in 155 countries. This means that, considering its existing opportunities and logistical infrastructures, Iran has been unable to effectively leverage its logistics infrastructures and economic opportunities to become a Middle East logistics centre. However, many researchers believe that this can and will be improved by 2025. Other neighbouring countries, on the other hand, benefited from the Middle East's peculiar situation in terms of logistics and economic growth in a proper manner, through proper understanding of environmental constraints, infrastructure improvements, and efficient management of their logistics systems. Emirates, Kuwait, and Turkey are among these nations, which rank ahead of Iran despite having less favourable geographical conditions. Iran's biggest competitors, Turkey, and the United Arab Emirates, were ranked 17th and 27th in the World Bank annual survey, respectively [25].

3. Methodology

This study in terms of method is considered quantitative research. A Likert type questionnaire tool was used while the statistical population of the research was the clients of 3PL companies in Tehran**.** From within these companies, executives or experts were recruited in the study to answer the questionnaire. These customers were from a variety of companies, from clothing to interior design and decoration. The sample size in this section was calculated by the Cochran formula:

$$N = \frac{Nt^2pq}{D^2(n-1) + t^2pq}$$

Where n is the sample size, N is the total number of statistical population, p is the presence of the trait, q is the absence of the trait, D is the acceptable error of 5% and t is the confidence level at 95% and is equal to (96/1). The number of clients of 3PL companies in Tehran follows an unlimited population, and based on the Cochran formula, 384 executives and experts were chosen from these client companies as the statistical sample.

This selection was done according to a random cluster model; different districts of the city of Tehran were considered as clusters and client companies were randomly selected from within these districts. The reason for using the random clustering method is that the intergroup variance is high regarding the statistical population, and by using this selection method an attempt was made to decrease this variance. The data after collection was analysed using SPSS software. First, the data was encoded and put into the software and then it was analysed using the two descriptive and inferential statistic methods. In the first and second sections, mean, variance, and frequency distribution tables and percentages were used to describe the opinions of the statistical sample regarding the questions. The significance level was set at 0.05. The results were analysed using descriptive and inferential statistics, Friedman's rank test, one-way t test, and mean comparison tests.

4. Data Analysis

The data analysis was performed after collecting the research data. The results related to the questionnaire of the 3PL companies' customers were considered to describe the barriers variable. Theses variables are the absence of precise review of the program, ambiguity in some items of the contract, high costs, legal barriers, management challenges, not doing the tasks properly, problems in the coordination of the company with the

3PL company, usage of non-specialist force, unpredicted cases, and wrong choice of the 3PL company in alphabetic order.

Table 2. The descriptive statistics of the questionnaire

	Numbers	Average	Standard deviation	Minimum	Maximum
Corporate legal barriers	384	2.94	1.16	1.00	5.0
Management challenges	384	3.25	1.25	1.00	5.0
Law barriers	384	3.38	1.16	1.00	5.0
Not doing the tasks properly	384	3.32	1.18	1.00	5.0
High costs	384	3.28	1.13	1.00	5.00
Use of non-specialist force	384	2.81	1.14	1.00	5.00
Wrong choice of the 3PL company	384	2.87	1.13	1.00	5.00
Absence of precise review of the programs	384	3.53	1.01	1.00	5.00
Problems in the coordination of the firm with the 3PL company	384	2.93	1.15	1.00	5.00
Ambiguity in some items of the contract	384	2.82	1.2	1.00	5.00
Unpredicted cases	384	3.07	0.86	1.00	5.00

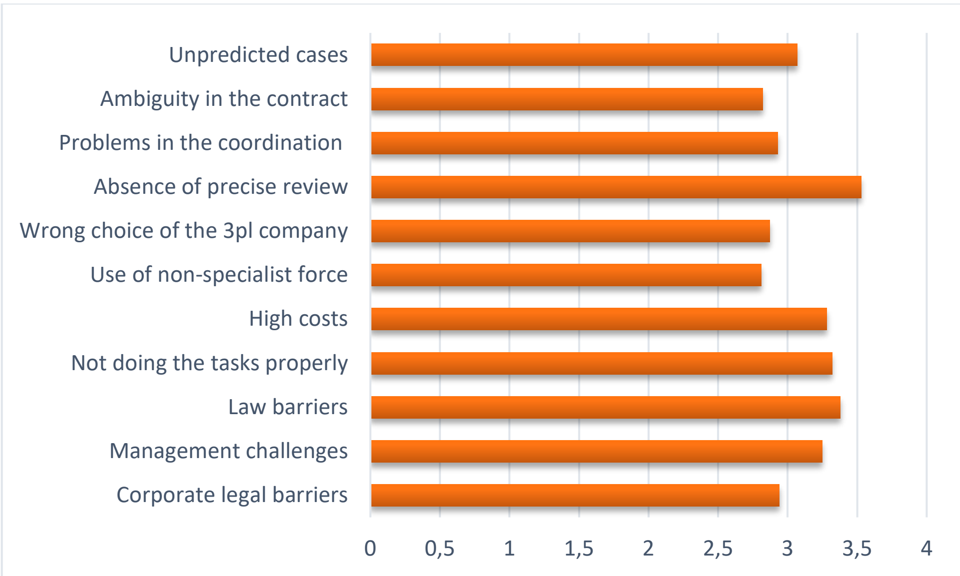


Figure 1: The descriptive statistics of the barrier's variable

According to the indices of the barriers, the factors related to absence of precise review of the programs has the highest average among the factors while factors indices related to management challenges, law issues, not doing the tasks properly and higher costs of outsourcing to 3pl companies seem to have higher scores among participants.

Table 3. The results of the one-sample T-test to identify the barriers

The society average = 3					
Variable name	Average difference	T-test statistic	Significance level	95% confidence interval	
				Lower bound	Upper bound
Corporate legal barriers	-0.06	-1.06	0.29	-0.18	0.05
Management challenges	0.25	3.93	0.001	0.12	0.38
Law barriers	0.38	6.37	0.001	0.26	0.49
Not doing the tasks properly	0.32	5.28	0.001	0.2	0.44
High costs	0.28	4.87	0.001	0.17	0.39
Use of non-specialist force	-0.19	-3.26	0.001	-0.3	-0.008
Wrong choice of the 3pl company	-0.13	-2.22	0.03	-0.24	-0.01
Absence of precise review of the programs	0.53	10.37	0.001	0.43	0.64
Problems in coordination of the company with the 3pl company	-0.07	-1.24	0.22	-0.19	0.04
Ambiguity in some items of the contract	-0.18	-2.93	0.001	-0.3	-0.06
Unpredicted cases	0.07	1.61	0.11	-0.02	0.16

To study the barriers, the index average was compared to the society average, which is equal to 3. The above average factors in table 1 were under T student test to check their significance and the results of Table 2 shows that in the intended model, the factors of management challenges, law barriers, not doing the tasks properly, high costs and absence of precise review of the programs are significant in the level of 0.05, while average of the factor unpredicted cases could not be trusted as significant. Comparing the average difference of this factor with the zero-mean difference of the society shows that these factors of the company are the main barriers for the retailers in outsourcing the activities of the 3pl companies.

5. Conclusion

Despite the unique logistical potential and advantages in Iran, so far for various reasons, the development of logistics and its infrastructure has been neglected in the country. The main challenges of the country's logistics field can be summarized as lack of a comprehensive document for the field of logistics, lack of multimodal transportation in the country, lack of formation and development of third-party logistics companies, lack of private sector participation in logistics development, disturbances and lack of coordination in the supply chain of commodities and related organizations, lack of development of logistics centres in Iran, the inability to change the country into a commercial hub in the region and the lack of physical infrastructure development.

The most significant limitations of retailers in Iran when outsourcing a logistics activity include management of challenges, official limitations, Improper completion of the task, high costs, and lack of expert knowledge of software programs. The most common reason is that logistics, as a basic function, is much more important than outsourcing itself; cost savings will not be made easily and integrating Iran's IT systems with the 3PL system is very difficult. While there is a particular logic behind each of these reasons not to outsource logistic responsibilities, solving some of these exact factors could accordingly encourage the related companies to farm out all or part of their supply chain responsibilities to 3PL potential partners. There is a great probability for the 3PL companies if they could possibly manage their legal issues to solve the related barriers, especially since it is noted as one of the most important barriers. The majority of 3PL companies' potential customers clearly indicated that failing to carry out tasks properly as well as the absence of precise review of the programs upset them. One could always claim that the lack of communication is the main reason for these barriers; implementing an adequate communication network and good presentation of "how outsourcing the logistic works" could potentially bring a market growth for 3PL companies. Based on this study, it is quite clear what the most significant barriers for the spread of 3PL in Iran are.

Iran is a country that connects the world's east and west, and it has long held a vital position in terms of trade, transportation, and logistics. However, there is currently only around 13,000 kilometres of road, which is insufficient for both passenger travel and foreign trade. Iran's airline industry is plagued by an out-of-date fleet. Lack of cooperation and communication with other modes of transportation, and poor airport profitability is also another major issue. The logistics industry in Iran is in desperate need of investment, and it is seen as a lucrative market for foreign investors. A wide market of 80 million people who are impacted by the current infrastructure (which is outdated, insecure, and slow) and a government which is monopolizing the logistic industry of the country by more than 60% and is not willing to compromise with the west to ease the sanctions.

References

[1] Alkhatib, S.F., Darlington, R., Yang, Z., Nguyen, T.T., 2015. A novel technique for evaluating and selecting logistics service providers based on the logistics resource view. Expert systems with applications, 42 (20), pp. 6976-6989.

[2] Alinejad, A., E., Pishvaee, M. S., & Bonyadi Naeini, A. (2018). Key success factors for logistics provider enterprises: an empirical investigation in Iran. Kybernetes, 47(3), pp. 426–440.

[3] Azad, N., Gh, Atiekar, 2016. The role of logistics and integrated supply chain management in industrial clusters to gain a competitive advantage. Case Study: Automotive Industry in Iran. International Conference of Management Elites. Tehran. Iran

[4] Berglund, M., Laarhoven, P.v., Sharman, G., Wandel, S., 1999. Third-Party Logistics: Is There a Future? The International Journal of Logistics Management, Volume 10, Number 1, 1999, pp. 59-70 (12)

[5] Boyson, S., Corsi, T., Dresner, M. and Rabinovich, E., 1999, Managing effective third-party logistics relationships: What does it take? Journal of Business Logistics, 20 (1), 73–100

[6] Bustinza, O.F., Arias-Aranda, D., Gutierrez-Gutierrez, L., 2010. Outsourcing, competitive capabilities and performance: an empirical study in service firms. International Journal of Production Economics, 126 (2), pp. 276-288

[7] Carter, C.R., 2005. Purchasing social responsibility and firm performance. International Journal of Physical Distribution & Logistics Management.

[8] Cheung, M.S., Myers, M.B., Mentzer, J.T., 2011. The value of relational learning in global buyer-supplier exchanges: a dyadic perspective and test of the pie-sharing premise. Strategic Management Journal, 32 (10), pp. 1061-1082.

[9] Choy, K., Chow, H., Tan, K., Chan, C., 2008, Leveraging the supply chain flexibility of third-party logistics - Hybrid knowledge-based system approach, Expert Systems with Applications 35 (4), pp. 1998-2016

[10] Cui, L., Gao, M., Sarkis, J., Lei, Z., & Kusi-Sarpong, S. In press. Modeling cross-border supply chain collaboration: the case of the Belt and Road Initiative. International Transactions in Operational Research. https://doi.org/10.1111/itor.12874

[11] Ebrahim-Khanjari, N., Hopp, W., Iravani, S.M., 2012. Trust and information sharing in supply chains. Production and Operations Management, 21 (3), pp. 444-464.

[12] Evangelista, P., Sweeney, E., 2006. Technology usage in the supply chain: The case of small 3PLs, The International Journal of Logistics Management 17 (1):55-74

[13] Evangelista, P., 2011. The role of information technology in supporting supply chain coordination of logistics service providers, in: Mahdavi, I., Mohebbi, S, Cho, N. (Eds), The book electronic supply network coordination in intelligent and dynamic environments: Modeling and implementation, IGI Global Publisher, pp. 113–144.

[14] Fang, J., Li, Q., He, Z.G., 2016. The solutions and recommendations for logistics problems in the collection of medical waste in China. Procedia Environmental Sciences, 31, pp. 447-456.

[15] Fu, S., Han, Z., Huo, B. 2017. Relational enablers of information sharing: Evidence from Chinese food supply chains. Industrial Management & Data Systems, 117 (5), 838–852.

[16] García-Morales, V.J., Jiménez-Barrionuevo, M.M., Gutiérrez-Gutiérrez, L., 2012. Transformational leadership influence on organizational performance through organizational learning and innovation. Journal of Business Research, 65 (7), pp. 1040-1050.

[17] Gigovic, L., Pamucar, D., Bajic, Z., Drobnjak, S., 2017. Application on GIS-interval rough AHP methodology for flood hazard mapping in urban areas. Water, 9 (6), 360 26 pages.

[18] Girod, S., Whittington, R., 2017, Reconfiguration, restructuring and firm performance: Dynamic capabilities and environmental dynamism, Strategic Management Journal, 38, (5).

[19] Ho, W., He, T., Lee, C., Emrouznejad, A., 2012. Strategic logistics outsourcing: An integrated QFD and fuzzy AHP approach, Operations & Information Management, Journal of Expert Systems with Applications.

[20] Huo, B., Ye, Y., & Zhao, X., 2015. The impacts of trust and contracts on opportunism in the 3PL industry: The moderating role of demand uncertainty. International Journal of Production Economics, 170160–170170.

[21] Jaberidoost, M., L, Olfat, A, Kebriaeezadeh. 2015. Pharmaceutical supply chain risk assessment in Iran using analytic hierarchy process (AHP) and simple additive weighting (SAW) methods. Journal of Pharmaceutical Policy and Practice, 8, Article number: 9, pp. 23-43.

[22] Lai, F., Tian, Y., & Huo, B., 2012. Relational governance and opportunism in logistics outsourcing relationships: Empirical evidence from China. International Journal of Production Research, 50 (9), pp. 2501–2514.

[23] Lieb, R. 2005. The use of third-party logistics services by large American manufacturers: The 2004 survey. Transportation Journal 44 (2), pp. 5-15.

[24] Lin, C., Wu, Y., 2014, Combined pricing and supply chain operations under price-dependent stochastic demand, Applied Mathematical Modelling, 38, (5–6).

[25] Kanwal, S., Pitafi, A.H., Ahmad, M., Khan, N.A., Ali, S.M. and Surahio, M.K., 2020. Cross-border analysis of China–Pakistan Economic Corridor development project and local residence quality of life. Journal of Public Affairs, 20 (2), pp. 2022.

[26] Mardani, A., Zavadskas, EK, Khalifah, Z., Nor, K., 2016, Multiple criteria decision-making techniques in transportation systems: A systematic review of the state-of-the-art literature, Transport 31 (3), pp. 359-385

[27] Morgan, R.M., Hunt, S.D., 1994. The commitment-trust theory of relationship marketing. Journal of Marketing, 58 (3), pp. 20-38.

[28] Navicelli, A. and De Carlo, F., 2018. Third-party logistics as a competitive advantage in Utilities spare parts management. IFAC-PapersOnLine, 51(11), pp.600-605.

[29] Pamucar, D., Mihajlovic, M., Obradovic, R., & Atanaskovic, P., 2017. Novel approach to group multi-criteria decision making based on interval rough numbers: Hybrid DEMATEL-ANP-MAIRCA model. Expert Systems with Applications, 88 (1), pp. 58–80.

[30] Panayides, P.M., 2007. Effects of organizational learning in third-party logistics. Journal of Business Logistics, 28 (2), pp. 133-158.

[31] Rahman, S., Ahsan, K.,Yang, L., Odgers, J., 2017. An Investigation into critical challenges for multinational third-party logistics providers operating in China. Journal of Business Research, http://dx.doi.org/10.1016/j.jbusres.2017.09.35.

[32] Santos-Vijande, M.L., López-Sánchez, J.Á., Trespalacios, J.A., 2012. How organizational learning affects a firm's flexibility, competitive strategy, and performance. Journal of Business Research, 65 (8), pp. 1079-1089.

[33] Selviaridis, K., Spring, M., 2007. Third party logistics: A literature review and research agenda. International Journal of Logistics Management. 18, pp. 125–150.

[34] Shi, Y., Zhang, A., Arthanari, T., Liu, Y., & Cheng, T. C. E., 2016. Third-party purchase: An empirical study of third-party logistics providers in China. International Journal of Production Economics, 171 (3), pp. 189–200.

[35] Vining, A., Globerman, S. 1999. A conceptual framework for understanding the outsourcing decision. European Management Journal, 17 (6), pp. 645–654.

[36] Vogel, R., Güttel, W.H., 2013. The dynamic capability view in strategic management: A bibliometric review. International Journal of Management Reviews, 15 (4), pp. 426-446.

[37] Wallenburg, C.M., Cahill, D.L., Michael Knemeyer, A., Goldsby, T.J., 2011. Commitment and trust as drivers of loyalty in logistics outsourcing relationships: cultural differences between the United States and Germany. Journal of Business Logistics, 32 (1), pp. 83-98.

[38] Wang, Y.F., Chen, S.P., Lee, Y.C., Tsai, C.T.S., 2013. Developing green management standards for restaurants: An application of green supply chain management. International Journal of Hospitality Management, 34, pp. 263-273.

[39] Yang, H., Zheng, Y. and Zaheer, A., 2015. Asymmetric learning capabilities and stock market returns. Academy of Management Journal, 58 (2), pp. 356-374.

[40] Zacharia, Z. G., Sanders, N.R., & Nix, N.W., 2011. The Emerging Role of the Third-Party Logistics Provider (3PL) as an Orchestrator. Journal of Business Logistics, 32 (1), pp. 40–54.

Biography

Daniel Sahebi (Born 1986), has been a Phd student and researcher since 2015. He has also been University Teacher in University of Vaasa in the School of Technology and Innovations since 2015 in the field of Logistics. Sahebi has been operating in the establishment of several start-up companies. He also is the chairman and the CEO of Oy Wasa CarWash Ab and Wasa CarService.

Petri Tapani Helo (Born 1975) has been Professor of Production Management, especially Logistics Systems, at the University of Vaasa since 2003. Helo operates in several technology companies. He is the Chairman of the Board of Wapice Oy , the Chairman of the Board of NoSwing Oy, a member of the Board of Jubic Oy and a deputy member of the Board of Merinova Oy .

2nd Conference on Production Systems and Logistics

Solution Approach For Challenges In Product Development In The Context of Emerging Cyber-Physical Systems

Henry Himmelstoß, Andreas Bildstein, Daniel Schel

Fraunhofer IPA, Stuttgart, Germany

Abstract

The emergence of cyber-physical systems (CPS), cyber-physical production systems and smart products is fundamentally changing the approach to product development. The paper addresses the current challenges in the development and design of product innovations arising from the emergence of cyber-physical systems, cyber-physical production systems and smart products, and the resulting opportunities. To provide a basic overview of the existing challenges, results from studies, scientific publications and experiences from industry are described. These results from industry are obtained from project work with a globally operating German manufacturing company. In the course of these challenges, companies must reorient themselves in terms of their products, services and find their role in changing ecosystems. The aim is to show the extent to these challenges addressed in science and the state of the art. One approach to address these challenges is described in this paper using the technical concept of the asset administration shell (AAS) and the CRISP-DM methodology. The goal is to provide insights in the early stages of product creation to drive incremental innovation or product improvement to facilitate the development of successful CPS and smart products.

Keywords

Product Innovation; Product Development; Industrie 4.0; Industry 4.0; Asset Administration Shell; Cyber-Physical System

1. Introduction

The digital transformation is permanently changing production environments and thus also production systems. The systems used in production are becoming increasingly intelligent and connected. As a result, in the final vision of Industry 4.0, production systems should be able to control themselves autonomously. To achieve this goal, cyber-physical systems and cyber-physical production systems represent a central enabler technology to enable new use cases. [1].

In turn, the emergence of cyber-physical systems is having a major impact on the way products are developed. Their development is characterized by high demands on interdisciplinarity, resulting in a new complexity in product development that must be mastered for the development of successful products [2].

In this context, cyber-physical products and cyber-physical production systems offer new opportunities. By connecting them to an infrastructure and enabling them to collect data end-to-end, new information can be obtained about their operation, maintenance or manufacturing. This is made possible by the ability of CPS to self-describe. Thus the CPS has knowledge about its own state or about its own capabilities; furthermore, the self-description can also contain information about the periphery or about the system context [3]. These information can in turn be applied in early phases of product development to design cyber-physical products in such a way that they can be sold more successfully on the market or meet the customer requirements in a better way. This requires new methods to select, process and make available this data for product development.

The aim of this paper is to address current challenges in the development of cyber-physical and smart products and to introduce a new approach to meet these challenges. Existing research is described and terminology work on cyber-physical systems and directly related technological concepts is presented in Chapter 2. Chapter 3 covers the current state of science as described by technical interest groups from Germany and in selected literature. In addition, findings from industrial work are considered. Chapter 4 describes an approach for a concept based on the asset administration shell and the Cross Industry Standard Process for Data Mining (CRISP-DM) procedure. Chapter 5 will briefly summarize the paper and conclude with an outlook.

2. Current state of research

This chapter explains the terms cyber-physical system and cyber-physical production system as well as the smart product based on them in context of this paper and the presented research approach. The Industry 4.0 component will also be described.

2.1 Cyber-physical product, cyber-physical production system and smart product

An early definition for cyber-physical systems dates back to 2006 [4]. Whereby CPS are described as integrations of computations with physical processes. Embedded computers and networks monitor and control physical processes, while physical processes affect computations due to feedback loops and vice versa [4]. The Association of German Engineers (VDI) defines CPS as a "system that links real (physical) objects and processes with information-processing (virtual) objects and processes via open, partly global information networks that are interconnected at all times". It also notes that a CPS can optionally use local or remote services and have a human-machine interface, and that the system can be adapted during its runtime. [5]. A cyber-physical production system (CPPS) is referred to when it is a CPS that is applied in a production-specific context [6].

Smart products are cyber-physical products that are enhanced by external infrastructure (internet) based smart services. A smart service is understood to be a service that is based on product-related data collected from the customer and that can be adapted to customer needs on this basis and can be used in a context-sensitive manner. Basically, it is stated that smart products generally consist of physical components (e.g. mechatronic components), intelligent components (e.g. software) and networking components (e.g. interfaces) [7]. Following this explanation, Figure 1 shows a simplified and generalized structure of the components of a smart product based on [7].

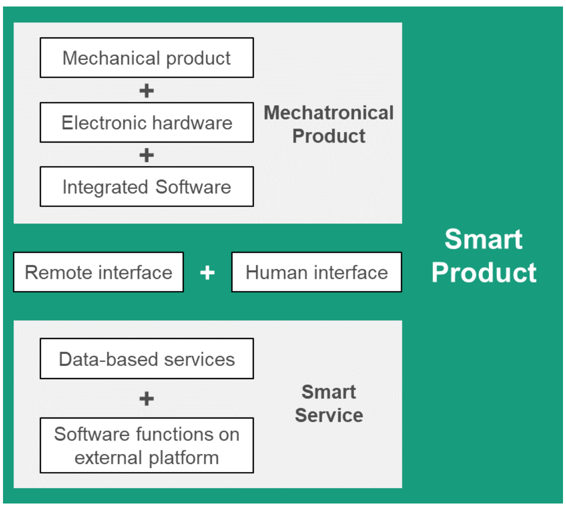


Figure 1: Simplified, generalized representation of the components of a smart product based on [7]

2.2 Industrie 4.0 Component

The combination of an asset with the AAS makes the Industrie 4.0 component. Every physical or logical object of value owned by a company or institution can be considered an asset. Capabilities can be attributed to each resource, therefore also to each asset. A capability describes an implementation-independent description of the functionality of a resource that has an effect in the real or virtual world and that serves for the fulfillment of a requirement for the processing a product. This means that the fulfillment of a required capability is checked against the capability of a resource to decide if a capability is the right match for the fulfillment of a desired task. Therefore before the execution of a certain task is initiated, a feasibility check is taking place to assess if a functionality is delivering the desired outcome. Capabilities can be in a hierarchy to each other and be composed of different capabilities [8]. The Industrie 4.0 component can therefore be regarded as the logical next evolutionary stage of a cyber-physical system in an industrial environment.

3. Current state in industry and academia

The development of CPS and smart products is complex and requires a high degree of interdisciplinary, since their components described in Chapter 2 require knowledge from different domains. At the same time, however, CPS offer new opportunities since they can allow data from the various phases of the product lifecycle to flow back into the early phases of product development, thus enabling a closed loop - field data from actual product usage can be used to derive new insights. This includes phases such as ideation, the concept phase and development [9].

This is supported by the concept that CPS are aware of their own state parameters as they pass through the different stages of their product lifecycle. Integrated into modern production flows, they can capture and store this data about themselves, as well as communicate and interact with other components in production systems [10]. However, a large part of the CPS currently in use mainly act passively, so that they primarily communicate clearly assignable information about their status in the form of status updates. Some of the CPS can also make or participate in decisions on their own. For example, information can be used to detect indicators of problems or to trigger defined actions [11]. The data generated by these products, for example through maintenance or repair, can be used to derive new products and strengthen customer relationships. This data can also be of great relevance when it comes to developing service-based capabilities for data processing [12]. Furthermore, with the help of CPS, companies can better capture information about actual user behavior, so that new insights will become apparent to companies about which product functionalities are actually used by product users and which receive little attention or are even avoided [13]. This can help companies establish a systematic approach to product improvement [14]. This is particularly important because the integrated development of physical products and digital services requires them to be developed in harmony with each other [15]. New tools are already being used to analyze the behavior of customers and consumers in order to offer products that precisely meet the desired characteristics. The development of products with the help of new technologies that take social and technological megatrends into account is already on the rise [16]. In a survey of experts this is underpinned by the fact that eight out of ten experts are of the opinion that information retrieval through smart products holds great added value for product development in that individual customer needs can be better addressed by new products [11].

Although the description of these possibilities takes place in scientific literature, concrete applications and procedures for the implementation in practice are currently still lacking. In particular, methods for identifying customer requirements are among the most important factors, but at the same time there is a lack of customer integration in the product development and product evaluation process [10]. This also corresponds to the findings of the Platform Industry 4.0 study on the engineering of smart products and services. In this study, research needs in various areas for the development of smart products were identified by analyzing relevant work and conducting interviews with experts. In this context, research needs are expressed several times when it comes to the use of product usage data for product development. For example, in the context of planning approaches, a need for research was mentioned for the provision of knowledge based on feedback information from previous product generations. In the context of requirements specifications and requirements management, for example, a research need was identified for approaches to generate knowledge from product usage data for new requirements specifications. According to the study, there is also a need for new approaches for the processing and for the provision of product utilization knowledge for engineering tasks [17]. It is also predicted that data from the product usage phase will increasingly be passed back into the development process and thus take on a more important role in the future [18].

During an industrial project with a component manufacturer who, like many representatives from the German SME sector, is facing the challenges of digital transformation, it was possible to identify needs similar to those in the study described earlier. Three significant topics that were identified during this work and are relevant for the considerations in this paper are summarized below:

- A concrete need for the collection of product data from the field has been identified. This data is to be utilized in particular with the aid of analytical methods for the improvement of existing products.
- Components that are already in use are able to collect data about themselves. However, there is uncertainty about how this data can be handled and used in the future. This goes hand in hand with the fact that it is currently still uncertain what role the collected data can play in future business models.
- Deriving a scope of services for future products proves to be difficult. Especially in the respect that it is unclear what functionalities will be demanded by customers in the future and expected in products they intend to buy.

From the identified needs of industrial partners, for which concepts are already described in the scientific literature, it is apparent that further solutions for industrial practice are needed.

4. Solution proposal

This chapter is describing the asset administration shell, the CRISP-DM and how these two concepts are being used in a new approach.

4.1 Asset Administration Shell

The AAS is a standardized representation of an asset and therefore a key factor for the interoperability between applications that are managing, steering and controlling manufacturing systems. The AAS and the asset that's represented by it can be uniquely identified - yet several AAS can be assigned to one asset, e.g. from the manufacturer's and the user's point of view. Submodels of different kind of aspects are contained and technical functionalities are described by the AAS [19]. Thus, by representing the characteristics of an asset through the AAS, it can be linked to the information world. These characteristics can be very wide-ranging and diverse. For example, they can be simple product properties, but also context-specific information about the product or information about its manufacturer. With an AAS containing multiple submodels, all information and functionalities of a given asset that a necessary for the completion of a dedicated use case can be described. Such information can include but are not limited to the features, characteristics, parameters, measurement data, capabilities or properties of the asset [8]. With the help of the AAS, a digital twin, i.e. a sufficient digital representation of an asset that meets the requirements of dedicated use cases, can be implemented [19]. These information can be covered by the AAS over the complete lifecycle of an asset and can therefore access its relevant data in every phase.

4.2 CRISP-DM

The CRISP-DM Methodology is described in a 6-step process model. It is a standardized procedure model that finds application in many data mining use cases. General tasks are assigned to each of these six phases, but are not discussed in detail further in this description. The process model consists of the phases business understanding, data understanding, data preparation, modeling, evaluation and deployment. However, the phases are not a rigid sequence, as it is often necessary to go through individual phases several times. In addition, there are other relationships between the phases that are not apparent in the general illustration in Figure 2. Which phase occurs next depends on the results achieved in each phase process model. There is a regular sequence, which also represents the main dependencies of the phases. This is also represented by the arrows between the phases. The large arrows arranged in a circle indicate the cycling nature of data mining, as it is not finished with the deployment, but re-introduces new findings. The CRISP-DM process model is shown in Figure 2 [20].

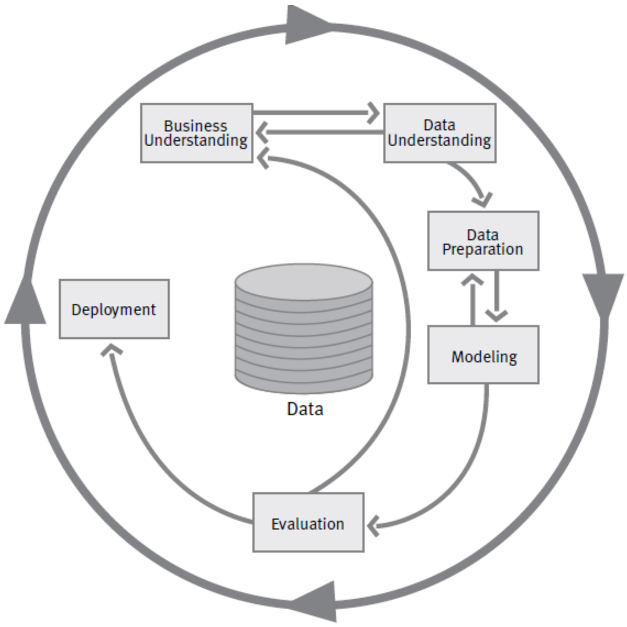


Figure 2: Phases of CRISP-DM [20]

CRISP-DM provides a very suitable framework for the methodological approach, since both aspects for evaluating use cases and technical aspects are given the necessary attention. Therefore, it is suitable for the methodological approach discussed in this paper, in that it also does justice to interdisciplinary problems and can be used in all domains.

4.3 Conceptual approach

Although it is described several times in the literature that product usage data can unfold new potentials in development, in practice there is still a lack of available solutions for industrial applications. Products and systems that have all the necessary characteristics of a smart product are increasingly being used in production environments. One approach to successful create new product adaptions and gain additional insights out of the usage of products can be to enrich product creation with additional knowledge from different phases of the product lifecycle. It also supports the product creation to cope with the new complexity in product development and to design products according to actual needs. Due to the given potential of the AAS to capture data of an asset along its product lifecycle, it is a suitable technological concept to realize use cases where data from the different phases of the product lifecycle are to be utilized in the early phases of product development. Figure 3 shows an exemplary product life cycle and the schematic feedback of data from the different phases into product development with the help of the AAS. In this context, assets are already linked to an AAS in their creation and carry information from this phase.

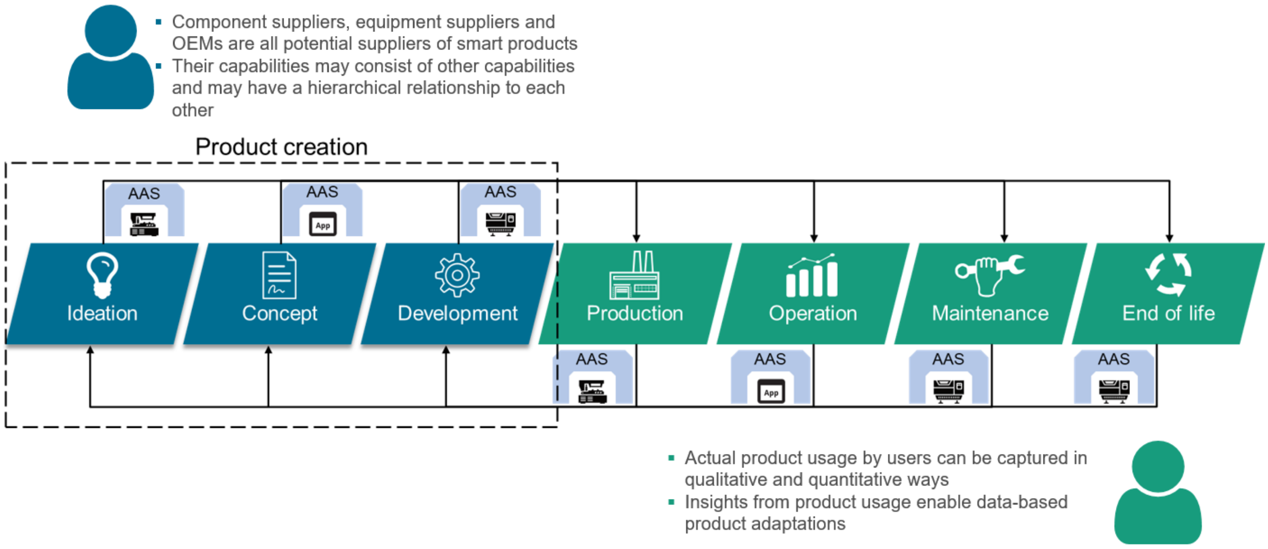


Figure 3: Exemplary product lifecycle with a closed feedback loop

In order to fundamentally assess the extent to which adaptations should be performed on a product, its quality in fulfilling specific tasks can be evaluated. This is to be based on the degree of fulfillment of a technology or a process, which results from the delta of the requirements of a product and the capabilities of a technology or a process. For this purpose, submodels of the AAS must be enriched with properties. These represent defined characteristics that are suitable for describing and differentiating products [19]. The source of this type of data continues to be the asset, with the AAS acting as a vehicle to enable a new type of interoperability and unique identification.

Such products or systems and the AAS create the technological foundation to enable closed feedback loops. To select and analyze the data from the product life cycle and for the early phases of product development, CRISP-DM is a suitable approach [20]. The combination of the AAS and the CRISP-DM procedure represents a novel research approach to systematically enrich product development with available knowledge from the product life cycle. However, since the data continues to come from the asset, it is necessary to consider which of its data is relevant and in what quantity and form it is required for the product adjustments.

	Business Understanding		Data Understanding
	Situation Assessment	I4.0 Component	Product Changes
Functionality based changes	▪ Description of requirements and associated formulation of assumptions ▪ Description of expected workflows ▪ Formulation of hypotheses for potential product improvements	▪ Definitions of submodels of the AAS with necessary properties ▪ Description of the capabilities ▪ Description of the hierarchy and composition of capabilities	▪ Initial Explorative recognition / detection of patterns in procedures ▪ Matching of recognized patterns with anticipated workflows and defined capabilities
Performance based changes			▪ Initial Collection of performance data according to previously defined capabilities ▪ Description of delta between target and actual capabilities
Business case based changes	▪ Description of requirements and associated formulation of assumptions ▪ Formulation of hypotheses for potential business case improvements		▪ Initial Collection of performance data & detection of patterns in procedures according to previously defined hypothesis ▪ Review of the current business model for weaknesses

Figure 4: AAS specific complementary tasks in the initial phases of the CRISP-DM

Therefore Figure 4 shows for the proposed approach necessary tasks complementary to the CRISP-DM in its early phases to derive product changes with the help of the AAS. A distinction is made between functionality-based, performance-based and business case based changes. While functionality-based product changes are derived from qualitative product usage, for example by not using functions or merging functions, performance-based product changes are derived from quantitative product usage, for example by using a product in a certain performance area in a defined period of time. Smart products also bring new challenges in the design of business models. Be it in the design of new monetization means, the design of value-based revenue models, or the use of product data in accordance with regulations. Therefore it is also important to continuously verify whether a smart product is contributing to the successful implementation and operation of a business model as planned. Therefore, it is necessary to continuously assess whether the currently operated business model meets the defined requirements and thus contributes to the planned business objective. The verification of the requirements and the associated hypotheses can be carried out with both qualitative and quantitative usage data.

For all types of product changes, the product requirements and related assumptions are described in the first phase of CRISP-DM. The assumptions will be used to formulate hypotheses about potential product improvements, which will be confirmed or falsified using real usage data. Massmann et al. propose a hypothesis-based and an explorative approach to data analysis. In the consideration of this thesis hypotheses are used for all types of product changes, however function-based product changes also require an explorative approach [21]. While anticipated usage in the hypotheses is described by workflows in the case of functionality-based product changes, anticipated target performance limits are described in the case of performance-based product changes. For potential adjustments to the business case, assumptions and hypotheses are continuously tested with real usage data to assess whether the intended business objective is being pursued in a target-oriented manner.

For all types of product changes, submodels of the AAS are defined whose properties and thus the collected data are suitable for testing the hypotheses. The capabilities of the Industrie 4.0 components are also described, what other capabilities they may be composed of, and what hierarchical relationship they have to each other.

This is followed by the second phase, Data Understanding. This phase involves an initial review and collection of data, as well as their general and qualitative description. In this phase, it is also necessary to identify which phases of the product life cycle are the source of relevant data and should therefore be the focus of the analysis. The first and second phases are particularly important for developing a concept for using the AAS and CRISP-DM to implement a closed feedback loop. However, the weighing of the need for product adjustments is conducted differently. For functionality-based product changes, patterns in product usage are exploratively sought and compared to anticipated workflows. A deviation of workflows from actual usage, for example, can provide insight into functions that are not necessary and thus can be eliminated if they are not used. However, new functions can also be derived from frequently recurring workflows in actual product usage - in this case, for example, a new, hierarchically superior capability can be derived from multiple lower-level capabilities. In the case of performance-based product change, threshold values formulated in hypotheses are compared with performance metrics from actual usage. This comparison is intended to evaluate the delta between the planned capabilities of a product and their actual manifestation in reality. It must also be taken into account whether the capability under consideration is a composite capability and what its hierarchical relationship is to others. The business case based adjustments can be supported by both the functionality-based and the performance-based usage data. For example, to price the operation of leased machines and systems in the high performance range with higher usage costs or to derive new offers from new functions.

Even if the focus is on these phases for the concept development, the other phases must also be addressed for a holistic assessment of the planned research. Accordingly, the data preparation and modeling phases must also address what data needs to be prepared and how, and what modeling techniques can help achieve the business objectives. In the evaluation phase, it is reviewed whether the results achieved meet the defined requirements or whether further iterations are needed.

5. Conclusion and Outlook

The use of CPS and smart products has also increased in industrial environments and the trend points towards increased development of products with intelligent services. Enriching product development with data from the field can be a facilitator in the development of successful products. Even though it is frequently described in the scientific literature that data from the field is fed back into development in a closed feedback loop, it is not often used in industrial applications so far. One of the reasons is that there are still hardly any solutions on the market that solve this problem. At this point, further research needs have been identified. In addition to technical applications, further conceptual work is also needed to systematically address such use cases in the future. In this paper, a new possible approach is presented. It is based on the technical concept of the AAS and the CRISP-DM procedure.

To verify this research approach, further conceptual elaboration is needed. In particular, a methodical approach is needed to make design decisions on the basis of parameters. It is also necessary to discuss the extent to which standardized submodels can be used for data-supported product development in order to proceed as standardized as possible. Also necessary is the application in industrial practice to put the approach to the test. In doing so, it is important to capture and sufficiently consider the needs and requirements of representatives from practice, so that a further step is taken for actual solutions to find their way into industrial practice, to enable closed feedback loops and to return product data from the life cycle to the early phases of product development. In this respect, further potential in product lifecycle management is also anticipated, as further valuable insights can be gained and provided for this through the interoperability of the AAS.

Acknowledgements

The research described in this paper is part of the Future Work Lab (FWL) - the innovation laboratory for work, people and technology of the Fraunhofer IPA and Fraunhofer IAO in Stuttgart. It represents a platform that combines numerous activities for researching the production work of the future. It also considers production systems, associated technologies and how these will be innovated in the future [22].

References

[1] Bauernhansl, T., Krüger, J., Reinhart, G., Schuh, G. WGP-Standpunkt Industrie 4.0. Wissenschaftliche Gesellschaft für Produktionstechnik WGP e.v. https://wgp.de/wp-content/uploads/WGP-Standpunkt_Industrie_4-0.pdf. Accessed 25 March 2021.

[2] Gausemeier, J., Dumitrescu, R., Steffen, D., Czaja, A., Wiederkehr, O., Tschirner, C., 2013. Systems Engineering in der industriellen Praxis. Heinz Nixdorf Institut, Universität Paderborn, Lehrstuhl für Produktentstehung | Fraunhofer-Institut für Produktionstechnologie IPT – Projektgruppe Entwurfstechnik Mechatronik | UNITY AG. https://www.hni.uni-paderborn.de/fileadmin/Fachgruppen/Seniorprofessur_Gausemeier/systemsengineerings/Studie_Systems-Engineering.pdf. Accessed 6 April 2021.

[3] Bauernhansl, T., Krüger, J., Reinhart, G., Schuh, G., 2016. WGP-Standpunkt Industrie 4.0. Wissenschaftliche Gesellschaft für Produktionstechnik WGP e.v. https://www.ipa.fraunhofer.de/content/dam/ipa/de/documents/Presse/Presseinformationen/2016/Juni/WGP_Standpunkt_Industrie_40.pdf. Accessed 6 April 2021.

[4] Lee, E.A., 2006. Cyber-Physical Systems - Are Computing Foundations Adequate?: Position Paper for NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap. Berkely EECS. https://ptolemy.berkeley.edu/publications/papers/06/CPSPositionPaper/Lee_CPS_PositionPaper.pdf. Accessed 11 April 2021.

[5] Verein Deutscher Ingenieure, 2020-09-00. Entwicklung cyber-physischer mechatronischer Systeme (CPMS). Beuth Verlag GmbH, Berlin. https://www.perinorm.com/document.aspx?hitnr=0&q=AC:DE45667245. Accessed 11 April 2021.

[6] VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik, 2017. Industrie 4.0 - Begriffe/Terms: VDI-Statusreport. https://www.vdi.de/ueber-uns/presse/publikationen/details/industrie-40-begriffeterms. Accessed 11 April 2021.

[7] Lünnemann, P., Wang, W.M., Lindow, K., 2019. Smart Industrial Products: Smarte Produkte und ihr Einfluss auf Geschäftsmodelle, Zusammenarbeit, Portfolios und Infrastrukturen. Fraunhofer IPA; CONTACT Software GmbH; VDI-Gesellschaft Produkt- und Prozessgestaltung (GPP). https://www.ipk.fraunhofer.de/content/dam/ipk/IPK_Hauptseite/dokumente/marktstudien/vpe-studie-smart-industrial-products.pdf. Accessed 13 April 2021.

[8] Bayha, A., Bock, J., Boss, B., Diedrich, C., Malakuti, S., 2020. Describing Capabilities of Industrie 4.0 Components. Federal Ministry for Economic Affairs and Energy (BMWi). https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Capabilities_Industrie40_Components.pdf?__blob=publicationFile&v=3. Accessed 25 March 2021.

[9] Rao, P., Palaniappan, K., 2020. Continuous Engineering Through ALM-PLM Integration, in: Product Lifecycle Management Enabling Smart X, Cham. 2020. Springer International Publishing, Cham, pp. 798–811.

[10] Nunes, M.L., Pereira, A.C., Alves, A.C., 2017. Smart products development approaches for Industry 4.0. Procedia Manufacturing 13, 1215–1222.

[11] Wuest, T., Schmidt, T., Wei, N.A., Romero, D., 2018. Towards (pro-)active intelligent products. IJPLM 11 (2), 154–189.

[12] Chowdhury, S., Haftor, D., Pashkevich, N., 2018. Smart Product-Service Systems (Smart PSS) in Industrial Firms: A Literature Review. Procedia CIRP 73, 26–31.

[13] Schuh, G., Zeller, V., Hicking, J., Bernardy, A., 2019. Introducing a methodology for smartification of products in manufacturing industry. Procedia CIRP 81, 228–233.

[14] Meyer, M., Frank, M., Massmann, M., Dumitrescu, R., 2020. Research and Consulting in Data-Driven Strategic Product Planning, In: Journal of Systemics, Cybernetics and Informatics, Vol. 18, Number 2, 2020.

[15] Meinhardt, S., Pflaum, A., 2019. Digitale Geschäftsmodelle - Band 2: Geschäftsmodell-Innovationen, Digitale Transformation, Digitale Plattformen, Internet der Dinge und Industrie 4. 0. Springer Vieweg. in Springer Fachmedien Wiesbaden GmbH, Wiesbaden.

[16] Miranda, J., Ponce, P., Molina, J.M., Molina, A., 2019. Taxonomy of Levels of Sensing, Smart and Sustainable Products to support New Product Development. IFAC-PapersOnLine 52 (13), 2384–2389.

[17] Abramovici, M., Gebus, P., Savarino, P., 2018. Engineering smarter Produkte und Services Plattform Industrie 4.0 Studie. acatech - Deutsche Akademie der Technikwissenschaften. https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/hm-2018-fb-smart.pdf?__blob=publicationFile&v=5. Accessed 25 March 2021.

[18] Abramovici, M., 2017. WiGeP-Positionspapier: „Smart Engineering“: Im Kontext von Industrie 4.0 werden sich Produkte und damit einhergehend auch das Engineering dieser Produkte radikal verändern. Wissenschaftliche Gesellschaft für Produktentwicklung WiGeP. http://www.wigep.de/fileadmin/Positions-_und_Impulspapiere/WiGeP-Positionspapier_Smart_Engineering.pdf. Accessed 25 March 2021.

[19] Sauer, M., Schaber, V., Schel, D., Schell, O., Schier, M., Schleipen, M., Schmitt, M., Schröder, T., Stojanovic, L., Teuscher, A., Uhl, A., Usländer, T., Vialkowitsch, J., Vollmar, F., Walloschke, T., Waser, B., Wende, J., Wiegand, M., Ziesche, C., 2020. Details of the Asset Administration Shell: Part 1 - The exchange of information between partners in the value chain of Industrie 4.0. Specification. Federal Ministry for Economic Affairs and Energy (BMWi). https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/Details_of_the_Asset_Administration_Shell_Part1_V3.pdf?__blob=publicationFile&v=5. Accessed 25 March 2021.

[20] 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. NCR Systems Engineering Copenhagen, DaimlerChrysler AG, SPSS Inc.,. https://www.the-modeling-agency.com/crisp-dm.pdf. Accessed 15 April 2021.

[21] Massmann, M., Meyer, M., Frank, M., Enzberg, S.v., Kühn, A., Dumitrescu, R., 2020. Framework for Data Analytics in Data-Driven Product Planning. Procedia Manufacturing 52, 350–355.

[22] Issa, A., Schumacher, S., Hatiboglu, B., Groß, E., Bauernhansl, T., 2018. Open Innovation in the Workplace: Future Work Lab as a Living Lab. Procedia CIRP 72, 629–634.

Biography

Henry Himmelstoß (*1991) is researcher at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA since 2019 where he is working in the department competence centre for Digital Tools in Manufacturing. In one of his roles, he is project manager of the innovation lab Future Work Lab.

Andreas Bildstein, Ph.D. (*1971) is head of 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.

Daniel Schel (*1986) is researcher at the competence centre for Digital Tools in Manufacturing at the Fraunhofer-Institute for Manufacturing Engineering and Automation IPA since 2011. He has managed and conducted a number of research projects as technical director and software architect, focusing on novel approaches to IT solutions in production.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Development of Digital Shadows for Production Control

Günther Schuh[1], Andreas Gützlaff[1], Matthias Schmidhuber[1], Judith Maibaum[1]
[1]*Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany*

Abstract

Production controllers have to make short-term decisions on the adjustment of production processes. Existing IT systems do not fully meet the information needs and functional requirements of production controllers and provide insufficient support for the decision-making process. The idea of Digital Shadows addresses this challenge by providing the user with the relevant information for a specific purpose consequently supporting the process of decision-making. This is achieved by linking, aggregating and analyzing production data from various sources.

In this paper, an approach for the development of Digital Shadows for production control is presented. The approach comprises the three levels of Digital Shadows for production control: 1) data offers, 2) information requirements and 3) analysis methods. To describe the data offers, the relevant data and data sources identified and categorized. The information requirements are derived from the decisions in production control. To link the data offers and information requirements the relations between the data and the tasks of production control are modeled. At the analysis methods level, existing production control analysis methods are examined and categorized. Additionally, an approach for the application of the analysis methods to the data is explored. The approach for the development of Digital Shadows was further on applied in order to create an exemplary Digital Shadow for production control.

Keywords

Digital Shadow; Production planning and control; Data analysis, Internet of Production

1. Introduction

Production control has the goal to realize the production planning under consideration of the dynamic production situation in the best possible way. [1] This significantly influences the performance of production and is therefore crucial for the success of a company. [2] However, production controllers face multiple challenges. A major challenge is a reaction to short-term changes and disruptions like sudden resource failures or deviations of the processing time from the plan lead time. The production controller must decide on appropriate changes to the production processes. [3,4] The effects of this decision's influence on the overall system can be hardly determined due to the high complexity caused by various interactions and dependencies in production. As a result, production control is often optimized only locally. [5]

Besides, external demands increase the performance requirements for production, which also affects production control. For example, customers require high adherence to delivery dates, short delivery times, and fast response to change requests which influences the logistic target values. Production control can influence the fulfilment of these demands, which requires a high flexibility of production control. [6,7] The demands for individual products and shortened product life cycles increase the complexity of production control, since the complexity of production control increases with the number of different orders and machines. [8,2]

At the same time Information Technology (IT) systems do not fulfil the increasing requirements of production control sufficiently. The IT systems do not contain all required information for a decision, why user manually create spreadsheets containing the relevant information. [9] Digital Shadows provide relevant information for a decision and thus represent a promising solution to overcome these challenges in production control.

Design and development of Digital Shadows is an objective of research within the Internet of Production, a German Cluster of Excellence at RWTH Aachen University. Digital Shadows collect, aggregate and analyse data from different sources to support the decision-making process. [6,10] For this purpose, the information required for a decision must be identified. [11] Information is generated by the application of analysis methods to aggregate, link and analyse the data. [12,13] Digital Shadows are a promising approach to improve a production system as they increase the decision-quality and provide users access to relevant information. [14]

The paper describes an approach for the development of Digital Shadows in production control. First, an overview of the theoretical background of Digital Shadows and Digital Twins is given. In section 3 related work on Digital Shadows for PPC is examined. Afterwards, the proposed concept for Digital Shadows in production control is illustrated. The application of the concept is explained in a use case. Finally, a conclusion on the current state is drawn with an outlook on future work.

2. Theoretical Background

In this section, the theoretical background on Digital Shadows is examined. This includes definitions and core elements as well as a comparison of both concepts.

A Digital Shadow consists of aggregated data traces and models. [15] The aim is to support the decision-making process. [14] For this reason, the right information must be provided in the right quality at the right place and at the right time. [6,16] For this purpose, the relevant information is generated within a Digital Shadow by linking, aggregating and analysing data traces using models. [17] A data trace consists of time-variant data and metadata. [12] A data-trace can originate from various IT-systems and are integrated by Digital Shadows. [18] The data-traces of a Digital Shadow describe both past and current states. [19] Changes in production create new real-time data traces. [15] Future states of the system can be derived from the application of simulation or prediction models. [19] As Digital Shadows are purpose-specific, different Digital Shadows exist. In production, they are used for e.g. production planning and order processing. [20]

NASA provided an initial definition of Digital Twins as "an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, etc., to mirror the life of its flying twin". [21] Current research often defines a Digital Twin as a virtual, complete representation of a physical system. [22,23,14,24] As a Digital Twin directly influences the physical system, a change of one system (e.g. virtual) leads to a change of the other system (e.g. physical). [25,23] Digital Twins analyse static and dynamic systems based on real-time production data. [26,19,24] A Digital Twin in production has the goal to optimize the production process by cost savings, efficiency increase and risk reduction. [25] In addition, a Digital Twin should support the decision-making process by analysing data and providing predictions. [27]

In summary, both Digital Shadows and Digital Twins support the decision-making process by providing information to the user. One difference is in granularity and scope. As a Digital Twin is a holistic representation of a system, it requires larger amounts of data and models than a Digital Shadow. A Digital

Shadow provides information for specific purposes, therefore only the necessary data and models are required.

3. Related Work on Digital Shadows in PPC

In this section, previous and related work concerning concepts for Digital Shadows and their use in the context of PPC are reviewed. Since Digital Shadows and Digital Twins contain similar elements, both approaches of Digital Shadows and Digital Twins for PPC are considered as related work. As of today, current research has barely addressed the design, use, and application of Digital Shadows for PPC.

Bauernhansel et al. [6] describe Digital Shadows for production as a macro-service providing the right information to the user. Various micro-services enable the functionalities of Digital Shadows. These are, for example, the collection of user requirements, the selection and compression of information and the control of the information flow. Four development stages of Digital Shadows are described: linkage of information, information flow control, information quality control as well as feedback and self-optimization of data and information basis.

The framework for Digital Shadows by Ladj et al. [28]consists of a virtual and physical system. The virtual part of a Digital Shadow is a data and knowledge management system. The knowledgebase is generated by the application of data analytic methods to the database. The information of the knowledgebase is used to support the decision-making process and to optimize the physical system. The Digital Shadow improves continuously, as it is self-learning. The application of the framework demonstrates improvements of the shop-floor performance.

Schuh et al. [29] model a data structure of Digital Shadows in the single and small batch production as a basis for a knowledge management system. The data structure is modelled based on a generic order fulfilment process in an Entity-Relationship-Model. The model visualizes the relationships between data and elements of the order fulfilment process. Considered data types are mater and transaction data of a production system based on the required information in the different departments of the order fulfilment.

Agostino et al. [30] describe a Digital Twin approach for PPC, which is based on a cyber-physical production system. The Digital Twin is a virtual representation of the production system. It consists of a simulation-based optimization model and functions triggering the simulation. The appliance of the results of the simulation optimizes the production system. Improved performance of production scheduling is demonstrated in a use case at an automotive supplier.

The approach by Schuh et al. [14] describes a top-down-bottom-up concept for designing Digital Shadows for PPC. Relevant information for the decision-making process in PPC is modelled top-down as information requirements. The information offers provided by analysing data with analysis models is modelled bottom-up. The concept helps companies designing Digital Shadows and applying them for PPC.

Kunath and Winkler's [5] conceptual framework of a decision support system for the order management process is based on a Digital Twin. The relevant data for order management is derived from different information systems. A suitable model regarding the data and decision is chosen by an automatic model generator. If no model is suitable a new simulation model is generated. Algorithms supporting the decision-making process evaluate the results of the simulation model. Possible applications of the framework are dynamic scheduling or dynamic pricing.

4. Digital Shadows for Production Control – Concept and Application

This section presents an approach for Digital Shadows for production control. First, a concept of Digital Shadows describing the elements and their relations is presented. Based on this, the three core elements data offers, information requirements, and analysis methods are specified by the application in a use case.

4.1 Concept of Digital Shadows for Production Control

The concept for Digital Shadows consists of three levels: data offers, information requirements, and analysis methods. By connecting the levels, Digital Shadows can provide the necessary information and thereby support the decision-making process. Figure 1 gives an overview of the concept of Digital Shadows. On the information level required information is determined. At the analysis methods level, the information requirements are transformed into data requests. The requested data is provided by the data offers level. The application of analysis methods on the data generates the required information and provides it to the users.

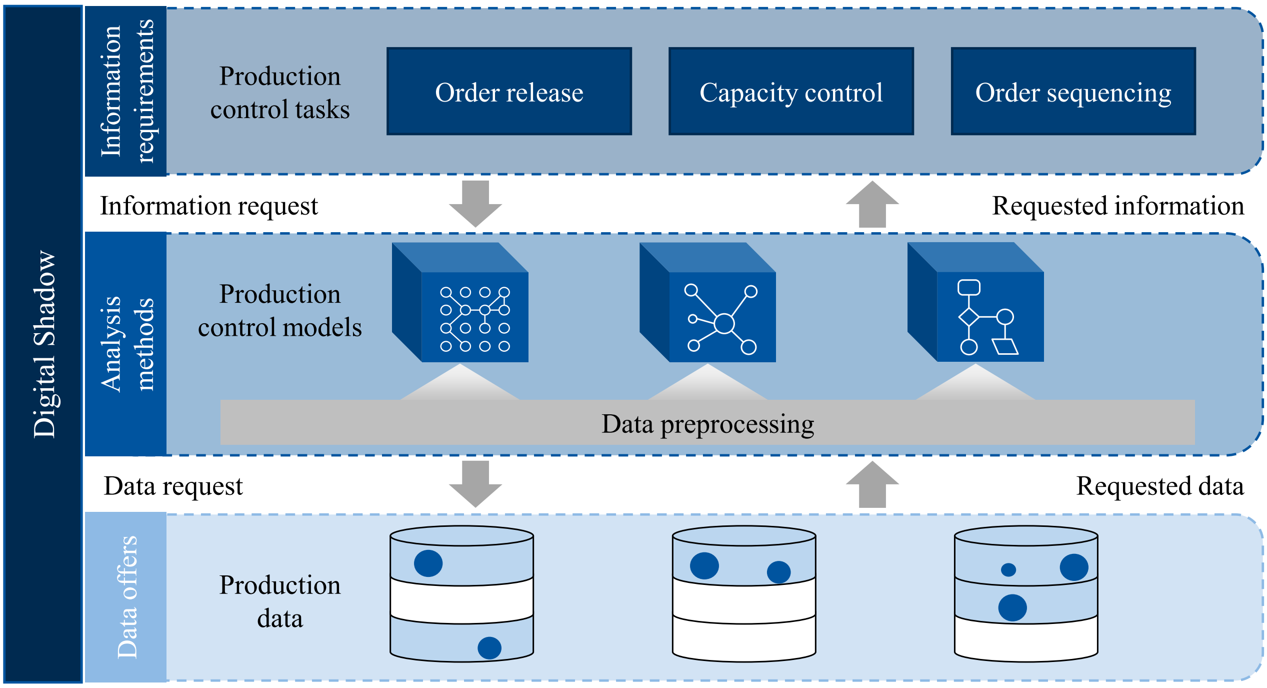


Figure 1 Concept of Digital Shadows for production control

The lowest level represents the data offers. On this level, data and relations between data are described. To model the data offers, elements of production relevant for production control are described and modelled. Relevant elements are e.g. orders, machines, personnel, and material. A Systems Modelling Language (SysML) based Block Definition Diagram (BDD) and Internal Block Diagram (IBD) can be used to give an overview of the elements and the relation of the elements. The production elements are the blocks of the model. Attributes describe the characteristics of the production elements. For production orders, attributes are e.g. delivery date, start date, product number and work schedule.

The data of each attribute is stored in the IT systems of a company. Both IT systems with a direct and indirect link to production control are possible data sources. IT systems with a direct link to production control are PPC systems, like Enterprise Resource Planning (ERP), Manufacturing Execution (ME), Advanced Planning and Scheduling (APS) systems. IT systems with an indirect link to production control are systems from other domains, such as development or purchasing, which contain relevant data for production control. These are e.g. Computer Aided Design (CAD), and Product Lifecycle Management (PLM) systems. As the IT system configuration is company-specific, the focus of this work is not on a generic description of the relationship between the IT system and data.

Data-specific features describe the data. These features contain information about e.g. source, type, quality, and generation frequency of data. The data offers are the foundation of Digital Shadows. They define the limits of the information generated by Digital Shadows in the current system configuration. If further information is needed to support the decision-making process, additional data must be collected and stored.

At the top level, the information requirements in production control are modelled. First, the decision-making requirements are derived from the tasks in production control. To realize the production plan, production control must initialize, monitor and adjust the production. The related tasks are e.g. order release, capacity control, and order sequencing. [31] Each of the tasks requires at least one decision from a production controller. For example, order release decisions determine the start and sequence of order release to the production. For each decision, the necessary information is determined and converted to information requirements. During the order release, information like delivery dates of open orders, and availability of resources and materials are required. Depending on the production control level (e.g. production network, unit, machine) different information is required.

The dimensions of information quality, type and quantity characterize the information requirements. Information quality describes the necessary accuracy, completeness and consistency. Information types range from descriptive, past-oriented types to predictive, future-oriented types. Information quantity describes the minimum viable data set to support the decision. [32] The modelled information requirements specify the information provided by the Digital Shadow through the model-based analysis of the data. The modelling of the information requirements thus represents the specifications for the digital shadow.

At the middle level, analysis methods connect the top and bottom level of Digital Shadows. Data is preprocessed and production control methods generate information. In this context, production control methods are mathematical models that generate information from existing data via e.g. simulations, optimizations or AI-based analyses. One example is the prognosis of the lead time based on master and transaction data like in [33–35].

Existing production control methods are classified based on their characteristics. Characteristics of the methods are input data, output information, and analysis type. The input data specifies the data type, necessary data quality and quantity, as well as the aggregation level of data. The output information specifies the information type, accuracy, and ease of interpretation. The analysis type describes the logic, e.g. simulation or neural network, the scope, e.g. prescriptive or predictive, as well as the properties, e.g. calculation duration, and traceability. By matching information needs and the data offers with the input and output variables of the production control models, suitable models for the specific decision supports can be identified.

Since the raw data from the system often does not meet the data quality requirements of the methods, the data is first preprocessed. This includes data integration, data cleaning, data normalization and data transformation. Data integration merges data of different sources to one data set. Data cleaning is necessary to correct and filter the wrong data. Data normalization and transformation adapt the convert the data into a form suitable for the production control model. [36] The Digital Shadow can then apply the models on the preprocessed data to generate the required information supporting the decision-making process.

4.2 Application of Digital Shadows for Production Control

The concept for Digital Shadows of production control described in section 4.1 is applied to a use case at the Demonstration Factory Aachen (DFA). DFA is a make-to-order manufacturer located in Aachen.

Since Digital Shadows are task-specific, an explicit Digital Shadow of production control is described. The focus of the use case is on decisions regarding short-term capacity adjustments in the context of capacity control. These are e.g. ordering of overtime as a reaction to disturbances like a resource failure. The production controller needs information about the configuration, costs and logistical target variables (adherence to delivery date, lead-time, capacity utilization and inventory [1]) of a future capacity scenario. The digital shadow must therefore provide future-oriented information on capacity scenarios in a second-by-second resolution. The observation period varies case-specifically from 1 day to 1 week.

The relevant elements of the system are the areas, such as logistics and production areas, machines, personnel, products, orders, material and IT systems. The IT systems are the source for the data of the individual attributes. Existing data sources are the ERP system as well as sensor data that collect electrical data. Figure 2 displays the elements as well as the descriptive attributes of the SysML BDD. The elements are production orders, products, machines, IT-Systems, personnel, material, transport vehicles and area. The attributes start-, end- and desired-date, work schedule, customer, order and product number, quality, and current state describe the production orders. The product number links every production order to one product. The attributes describing a product are product group, BOM, and work plan. The work plan defines which capacities are required to fulfil a production order. The work schedule defines when the capacities are used. Personnel and machines are the capacities to fulfil the production orders. Machine group (e.g. laser cutting, saw), hourly wage, number of workplaces, shift, wait-time, lay-time as well as electrical power and currency describe the machines. Personal group (e.g. assembly, welding), qualification, shift and hourly wage characterize the personnel. Attributes characterizing the material are material type, costs, number, stock location, quantity in stock, supplier and supply time. The BOM links each production order to the required material. Transport vehicles transport materials and products within the DFA. The vehicle type (e.g. forklift, electric crane), transport speed and quantity as well as the availability describe them. Relevant attributes for areas are size, required personnel and area type like storage, production or assembly. Either the ERP-system or sensors provide the data. For the task of capacity control, not all modelled data is necessary. Only the data used by the analysis methods are required.

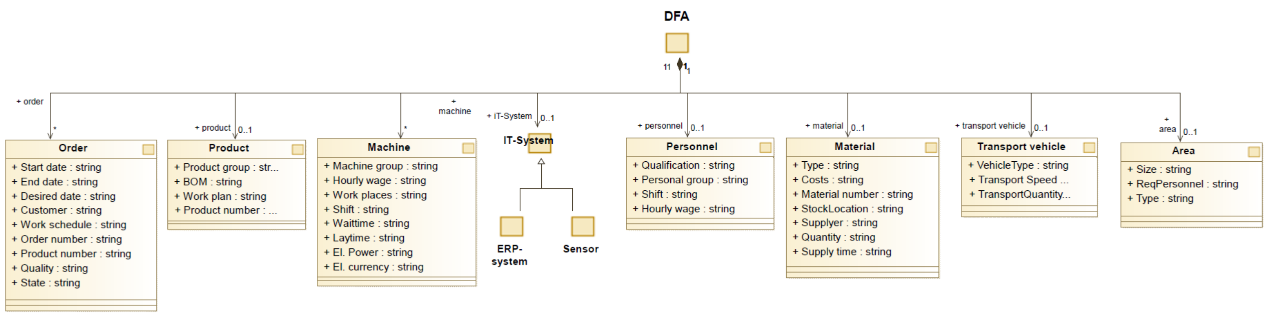


Figure 2 Block Definition Diagram of elements and attributes of DFA

To fulfil the user's information requirements, suitable analysis methods must be identified. To support the decisions of capacity control the Digital Shadow must calculate and evaluate different capacity plans. As a simulation can calculate and evaluate different production scenarios, it fits the requirements best. It is used as an analysis method for this use case. The simulation determines the occupancy of machines and personnel by orders in the simulation period under consideration of material availability (here between 1 day and 1 week). Required data are the available capacities of machines and personnel as well as the required capacities of production orders. The possible processing time for a production order is limited by material availability and the planned start and finish dates. A calculated capacity plan is then evaluated regarding costs and adherence to logistic targets. Data of machine and personnel costs are required. Table 1 gives an overview of all required data for the simulation and the related attributes of the BDD.

The simulation calculates and evaluates different scenarios with different capacity configurations. The results are information on possible capacity plans with their costs and logistical target values. The Digital Shadow provides this information to the user supporting the decision-making process of capacity controlling.

Table 1 Required input data for simulation model

Production element	Required data for simulation model	Attributes in BBD
Machine	Available capacities	Shift plan
	Machine costs	Hourly wage
Personnel	Available capacities	Shift plan
	Personnel costs	Hourly wage
Order	Possible processing period	Planned start date
	Possible processing period	Desired end date
	Allocation to product	Product number
Product	Required capacities	Work Plan
	Required material	BOM
Material	Availability material	Availability

In summary, the Digital Shadow enables production controllers to evaluate different capacity scenarios and decide on the best scenario according to the current targets. This increases transparency and improves the target achievement.

5. Summary and Outlook

The dynamic production environment as well as increased demands in production control require an improved decision-making process support for production controllers. The concept of Digital Shadows is a promising approach to provide the user with relevant information. In this paper, an approach for developing Digital Shadows for production control is presented. The three levels of the proposed concept are data offers, information requirements and analysis methods. The analysis methods link the data offers with the information requirements by generation the information based on the offered data. The concept is applied to a real-world use case at DFA. The developed Digital Shadow supports decisions of capacity control.

In comparison to existing IT systems for PPC, the Digital Shadow provides all required information for a decision in production control. Data of different systems and domains are aggregated to the required detail level. Additionally, Digital Shadows use task-specific, in-depth analysis methods generating detailed information. In conclusion, the Digital Shadow provides the relevant information for a decision improving production control.

The presented approach focuses on one use case. In further research, more use cases and the application of the concept at different companies can be addressed. In the actual configuration of the Digital Shadow, users choose the scenarios. In future realizations, an AI can be integrated to generate scenarios. In addition, the linkage of different production control models to generate information can be further examined.

Acknowledgements

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

References

[1] Schuh, G., 2007. Produktionsplanung und -steuerung: Grundlagen, Gestaltung Und Konzepte. Springer, Dordrecht, 887 pp.

[2] Frazzon, E., Kück, M., Freitag, M., 2018. Data-driven production control for complex and dynamic manufacturing systems. CIRP Annals - Manufacturing Technology 67 (1), 515–518.

[3] Steinlein, F. Decision support app for short term production control to improve the adherence to delivery dates.

[4] Yang, S., Arndt, T., Lanza, G., 2016. A Flexible Simulation Support for Production Planning and Control in Small and Medium Enterprises. Procedia CIRP 56, 389–394.

[5] 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.

[6] Bauernhansl, T., Hartleif, S., Felix, T., 2018. The Digital Shadow of production – A concept for the effective and efficient information supply in dynamic industrial environments. Procedia CIRP 72, 69–74.

[7] Hauptvogel, A., 2015. Bewertung und Gestaltung von cyber-physischer Feinplanung. Dissertation, Aachen, 246 pp.

[8] Ellwein, C., Elser, A., Riedel, O., 2019. Production planning and control systems – a new software architecture Connectivity in target. Procedia CIRP 79, 361–366.

[9] 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 Produktioinsplanung und -steuerung: Ergebnisse einer gemeinsamen Studie der produktionstechnischen Institute Fraunhofer IGCV, IFA, IPMT und WZL. Zeitschrift für wirtschaftlichen Fabrikbetrieb ZWF 115 (4), 252–256.

[10] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., Sui, F., 2018. Digital twin-driven product design, manufacturing and service with big data. The International Journal of Advanced Manufacturing Technology 94 (9-12), 3563–3576.

[11] Riesener, M., Schuh, G., Dölle, C., Tönnes, C., 2019. The Digital Shadow as Enabler for Data Analytics in Product Life Cycle Managemen. Procedia CIRP 26, 729–734.

[12] Schuh, G., Häfner, C., Hopmann, C., Rumpe, B., Brockmann, M., Wortmann, A., Maibaum, J., Dalibor, M., Bibow, P., Sapel, P., Kröger, M., 2020. Effizienter Produktion mit Digitalen Schatten. ZWF Sonder, 729–734.

[13] Schuh, G., Prote, J.-P., Gützlaff, A., Thomas, K., Sauermann, F., Rodemann, N., 2019. Internet of Production, in: Wulfsberg, J., Hintze, W., Behrens, B. (Eds.), Production at the leading edge of technology. Proceedings of the 9th congress of the German Academic Association for Production Technology (WGP), 1 ed. Springer Vieweg Berlin, Heidelberg, pp. 533–542.

[14] 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.

[15] Brecher, C., Buchsbaum, M., Storms, S., 2019. Control from the Cloud: Edge Computing, Services and Digital Shadow for Automation Technologies. International Conference on Robotics and Automation, 9327–9333.

[16] Pause, D., Brauner, P., Faber, M., Fischer, M., Hunnekes, P., Petruck, H., Mertens, A., Nitsch, V., Schuh, G., Stich, V., Ziefle, M., 2019. Task-Specific Decision Support Systems in Multi-Level Production Systems based on the digital shadow. IEEE International Conference on Industrial Engineering & Engineering Management, 603–608.

[17] Riesener, M., Dölle, C., Schuh, G., Tönnes, C., 2019. Framework for defining information quality based on data attributes within the digital shadow using LDA. Procedia CIRP 83, 304–310.

[18] Schuh, G., Dölle, C., Tönnes, C., 2018. Methodology for the derivation of a digital shadow for engineering management. IEEE International Conference on Industrial Engineering & Engineering Management, 1–6.

[19] Stecken, J., Ebel, M., Bartelt, M., Poeppelbuss, J., Kuhlenkötter, B., 2019. Digital Shadow Platform as an Innovative Business Model. Procedia CIRP 83, 204–209.

[20] Kubenke, J., Roh, P., Kunz, A., 2018. Assessing the Efficiency of Information Retrieval from the Digital Shadow at the Shop Floor using IT Assistive Systems, in: Reinventing Mechatronics : Proceedings of Mechatronics 2018. 16th Mechatronics Forum International Conference, Glasgow, United Kingdom. 19.-21. September 2018, pp. 202–209.

[21] Shafto, M., Conroy, M., Boyle, R., Glaessgen, E., Kemp, C., LeMoigne, J., Wang, L., 2012. Medeling, Simulation, Information Technology & Processing Roadmap. Technology Area 11, 1–38.

[22] 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.

[23] 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.

[24] 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. International Conference on Process Control 21, 258–262.

[25] 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.

[26] 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.

[27] Qi, Q., Tao, F., 2018. Digital Twin and Big Data Towards Smart Manufacturing and Industry 4.0: 360 Degree Comparison. IEEE Access 6, 3585–3593.

[28] Ladj, A., Wang, Z., Meski, O., Belkadi, F., Ritou, M., Da Cunha, C., 2021. A knowledge-based Digital Shadow for machining industry in a Digital Twin perspective. Journal of Manufacturing Systems 58, 168–179.

[29] Schuh, G., Kelzenberg, C., Wiese, J., Ochel, T., 2019. Data Structure of the Digital Shadow for Systematic Knowledge Management Systems in Single and Small Batch Production. Procedia CIRP 84, 1094–1100.

[30] Agostino, Í.R.S., Broda, E., Frazzon, E.M., Freitag, M., 2020. Using a Digital Twin for Production Planning and Control in Industry 4.0, in: Sokolov, B., Ivanov, D., Dolgui, A. (Eds.), Scheduling in Industry 4.0 and Cloud Manufacturing, vol. 289, 1st ed. 2020 ed. Springer International Publishing; Imprint: Springer, Cham, pp. 39–60.

[31] Lödding, H., 2008. Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration, 2., erw. Aufl. ed. Springer Berlin Heidelberg; Springer-Verlag, Berlin, Heidelberg, 579 pp.

[32] Hempel, T., 2018. Bedarfsgerechte Informationsbereitstellung zur hochauflösenden Produktionsplanung und -regelung. Dissertation, Aachen, 244 pp.

[33] Gyulai, D., Pfeiffer, A., Bergmann, J., Gallina, V., 2018. Online lead time prediction supporting situation-aware production control. Procedia CIRP 78, 190–195.

[34] Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., Monostori, L., 2018. Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. Procedia CIRP 72, 1051–1056.

[35] Schuh, G., Prote, J.-P., Sauermann, F., Franzkoch, B., 2019. Databased prediction of order-specific transition times. CIRP Annals 68 (1), 467–470.

[36] García, S., Luengo, J., Herrera, F., 2015. Data Preprocessing in Data Mining. Springer International Publishing; Springer, Cham, 327 pp.

Biography

Günther Schuh (*1958) holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University. Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh is a 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. He serves on several supervisory and administrative boards.

Andreas Gützlaff (*1989) is research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and is head of the Production Management department since 2019. He studied industrial engineering with a specialization in mechanical engineering at RWTH Aachen University.

Matthias Schmidhuber (*1992) is research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University and is head of the Production Logistics group in the Production Management department. since 2017. He studied industrial engineering with a specialization in materials and process technology at RWTH and at Aalto University in Helsinki

Judith Maibaum (*1994) is research assistant at Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University in the Production Logistics group of the Production Management department since 2018. She studied industrial engineering with a specialization in mechanical engineering at RWTH Aachen University.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Requirements for Human-Centric Informational Complexity Management in Production in the Context of the Matrix Fusion Factory

Jörg Siegert[1,2], Thilo Schlegel[1,2], Liliana Zarco[1], Thomas Bauernhansl[1,2]
[1]Institute of Industrial Manufacturing and Management IFF, University of Stuttgart, Nobelstr. 12, Stuttgart, 70569, Germany
[2]Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstr. 12, Stuttgart, 70569, Germany

Abstract

For four years now, the concept of the Matrix Fusion Factory MFF has been developed and studied. The MFF dissolves the separation between digital and real entities and focuses on value creation in production. The MFF is based on two prerequisites: first, an organizational structure for assessing value creation in factories whose modular hardware and software structures are constantly reconfigured; and second, the fragmentation logic presupposes increasing complexity, which in turn requires methods for evaluating and reducing complexity. This paper focuses on the reduction and management of informational complexity since different ways of information management (i.e. capturing, selecting, compressing, and providing information) lead to different effects on the production system, especially if the information is processed by humans. Therefore, the central assessment approach for information management must be value creation. Thus, this paper discusses the impact of information management on value creation and how information management can support value creation. In particular, it is clarified how informational complexity in production can be reduced without distorting the underlying information and the implications for technology and organization are discussed.

Keywords

Matrix Production; Complexity Management; Information Management; Matrix Fusion Factory; Decision Making; Information Logistics

1. Introduction and Motivation

Ultra-flexible, mobile, and highly modularized manufacturing systems consisting of constantly reconfigured hardware and software module networks offer numerous additional degrees of freedom due to their modularity and mobility. To enable optimal value creation, these degrees of freedom require methods and approaches to deal with the enormous increase in complexity. [1–9]

The basic requirement for creating value in production is to complete customer orders with assured quality. Especially in ultra-flexible factories, where the entire manufacturing system is in a constant state of change, the order becomes the only constant that cannot be changed by the manufacturing system [2, 10–13]. Therefore, the MFF considers the factory as a service provider for the respective order. The goal is to optimally adapt the manufacturing system to the order.

The completion of the customer's order is the basis of value creation, with the fundamental objective being maximum added value through a minimum use of resources. The MFF therefore aims at minimizing the use of parts and factory resources to complete an order. Prerequisites for the minimum use of production resources are scalability of the resources themselves and sufficient degrees of freedom to adapt to the order. The MFF considers machines as context-related, temporary networks of scalable hardware and software modules, thus generating additional degrees of freedom [2].

According to Milling, a system is complex if the number of possible links within a system is no longer manageable and the causalities among them cannot be recognised anymore [14]. According to Kluth et al. complexity involves the four complexity dimensions variety, heterogeneity, dynamics, and non-transparency [3]. According to Siegert et al., complexity in modular and flexible production systems depends on the number of degrees of freedom of its subsystems and components [1]. These degrees of freedom include, above all, the spatial, time and informational forms, interconnections and changes of all entities of a system. However, even if complexity may not be quantified, the varying states of complexity of a production system can be described and differentiated from each other [1]. According to Weaver complexity can be -classified as organised and unorganised complexity [15]. Ferretti defines a system's complexity as ontological complexity, the existing lack of order, or as epistemological complexity, i.e. the overstraining of human perceptual capacities because of a wide variety and diversity of existing interdependencies [16].

Both extreme ontological and (perceived) epistemological complexity lead to an overload of information. Therefore, the concept of minimizing the use of resources is also applicable to information. In line with the other production resources, it must be decided how much and which information is required to fulfil a task. The effect of distorted information must therefore be considered.

Each decision and each process in a manufacturing system is based on information [17–19]. In particular, the above-mentioned degrees of freedom can only be used to add value by modelling information [20]. Value creation is therefore dependent on the right information in the right place at the right time [21, 22]. For this reason, information, its provision, and its processing require special consideration in the MFF.

The underlying research questions are therefore:

- What is the role of information and information distortion concerning value creation in production? Especially considering the concept of a minimum use of production resources?
- How to determine the admissible value/degree of information
- distortion?
- What requirements does this place on organizational and technical structures? How can these requirements be addressed?

This paper uses the paradigm of the Matrix Fusion Factory to show how value creation can be optimized by incorporating relevant production aspects. In the following, the role of information and information distortion in the concept of a minimum use of resources is described and how to determine the permissible degree of information distortion. Then, the role of the decision maker is considered from a human and machine perspective before explaining the organizational structures of the MFF.

2. Determining permissible information distortion

Distorting information means selecting, presenting, or weighting information in such a way that the underlying facts described by the information are framed in a certain way. This usually results from omitting or not considering some additional information by oneself or others. That does not necessarily mean reducing the content of truth of the various pieces of information per se or creating false information. Thus, it is not necessarily the information itself that is distorted, but the picture that is gained from the information [23–28]. Information distortion is therefore a cognitive bias [25–28] which in extreme cases may also lead to false knowledge.

Any way of aggregating information therefore leads to a distortion of the underlying facts [27]. However, since humans [29] or computers are unable to solve tasks unless information is aggregated, this is an essential step [30]. Therefore, it is not necessary to clarify whether information may be distorted but rather to which extent this is permissible.

This poses the problem of how to identify potentially distorted information. Checking the content of truth of individual pieces of information is certainly not the sole solution because the truth of individual pieces of information may be unchanged. It is only possible to detect the absence or incorrect weighting of individual pieces of information causing a distortion, if the overall picture generated can be matched some other way. This can be achieved by comparing it either with the real facts or with other images with validated truth content.

The relevant production target variable is quality-assured added value. In production, information is usually condensed to aid decision making. For this purpose, complex facts are simplified, or important interrelationships are highlighted. Whether the impression of a fact has been "mapped correctly" can therefore be determined by the impact of the decision made. If the decision or the resulting action adds value, the distortion of information was permissible or at least not detrimental. However, if the decision has a negative effect on value creation, the distortion was not permissible. The basis for assessing permissible information distortion is therefore its effect on value creation.

This leads to a further examination of the decision-making process, its influences and the resulting actions within the production system.

Understanding mental processes, such as perception, memory, reasoning, and motor response, as they affect interactions among humans and other elements of a system in general, is an independent discipline called cognitive ergonomics [31].

According to Endsley, in cognitive ergonomics decision making and the subsequent performance of the decided course of action can be traced down to the situation awareness of the decision-maker [29, 31, 32]. Endsley [33] defines situation awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". Dominguez defines situation awareness as the "continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events" [34].

Consequently Endsley's model [29, 32] (see Figure 1) is set into three levels of situation assessment, each level being a necessary precursor to the next higher level. This model follows a chain of information processing, from perception, through interpretation, to prediction. The three levels of situation awareness are as follows: Perception of the elements in the environment: this stage is the first step to achieve the situation awareness which is related to the human's perception of information in a given time and space. Comprehension of the current situation: comprehension is essential to understand the significance of the elements perceived and to gain a

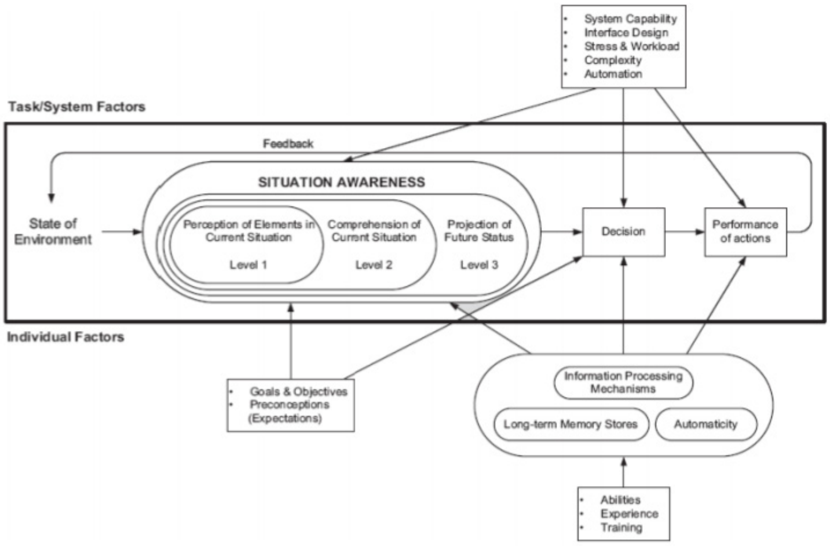


Figure 1 Model of situation awareness levels [29]

picture of their relationships. Projection of future status: this level of situation awareness is associated with the ability to project the future of the elements in the environment. Accuracy of the projection is highly dependent upon the accuracy of perception and comprehension.

In the context of decision making, information distortion can therefore be understood as the wrong perception of information or the wrong supply of information, or the wrong comprehension of the current situation and therefore a wrong picture of the relationships of the perceived elements. This will lead to a wrong projection of a future status and subsequently to a poor decision and taking a suboptimal decision.

According to Endsley's model, the parameters influencing the situation awareness and therefore inducing information distortion can be divided into individual factors and task/system factors.

Individual factors can be influenced by training and experience. This is underlined by the conclusion of Carley and Lin that information distortion can at least to some extent be combated by training [27] and ties in with IFF's definition of competence [17]: "Competencies represent the entirety of abilities, skills and existing knowledge to adequately apprehend, analyse, evaluate, make decisions, and act correctly in complex, dynamic and sometimes chaotic situations, taking into account relevant goals." This definition strongly emphasises the evaluation of the situation to take the correct action. Therefore, systematically correct action calls for the analysis and evaluation of the situation and for decision-making. Unsystematically correct action, i.e., action that is not in line with the decision but nevertheless adds value, makes the system unpredictable in the long term if it occurs frequently and should therefore be avoided. If unsystematically correct action is taken frequently, decision making processes or decisive entities should be checked for their suitability.

Task/System factors are largely influenced by complexity and the resulting stress and mental workload. Mental workload reflects the amount of mental resources required to perform a set of concurrent tasks [35]. Manufacturing systems are therefore prone to diminishing the situation awareness and to inducing information distortion of the individuals since they typically have a high level of complexity, many different capabilities, are highly automated, and are usually operating under cost and time pressure, therefore creating a high stress and high workload environment for their employees to work in. This impedes the situation awareness of the decision makers and leads to information distortion. This is underlined by the findings of Chaxel, Wiggins et al., that the mere belief of having limited time to make a decision is influencing and aggregating information distortion [23]. In addition, according to Carley and Lin, technology based distortion is typically more debilitating than personnel induced information distortion [27]. Especially as far as big data approaches are concerned, this leads to the consideration of how to reduce the content of the various dimensions of the 5V model [36] without unduly distorting it.

In short, to ensure correct actions and therefore value creation, the individual factors and the task/system factors must thus be brought into balance (see Figure 2). This paper focuses on the organizational and technical requirements for adjusting the task and system factors to the individual factors to avoid information distortion.

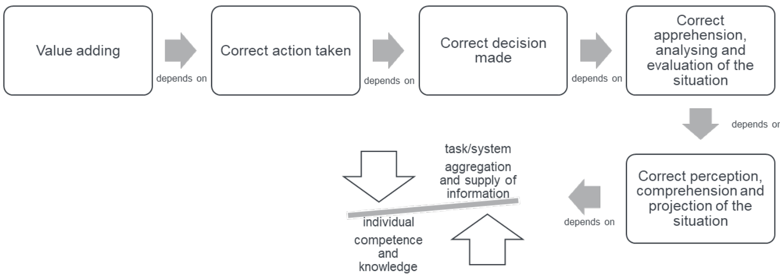


Figure 2 Value adding and balancing factors

3. Determining optimal degrees of freedom

In manufacturing systems, decisions are made by different actors or entities. Decisions can therefore be divided into decisions made by machines and decisions made by humans, depending on the type of entity making the decision. On a higher level, the process leading to the making of a complex decision is similar both for humans and machines, regardless of their different data processing capacities and speeds. Both need to consider the facts of the matter on the basis of information, weigh this information, determine possible courses of action, prioritize these courses according to the anticipated effects, and then decide on one of the courses of action [37, 38]. The fundamental difference between decisions made by machines and by humans apart from the suitability for different situations and areas of application is therefore from the industrial management point of view the assumption of responsibility [17].

Greater degrees of freedom in the manufacturing system mean more possible courses of action and more scopes for action. In general, this means that the number of suboptimal courses and scopes for action increases disproportionately as the degrees of freedom increase. This is problematic because the possibility of making wrong or suboptimal decisions may not only have a negative effect on value creation but may also unnecessarily increase epistemological complexity. The decision maker (whether human or machine) is confronted with a multitude of decision options, from which the few relevant options must first be laboriously filtered out before the most suitable one can be determined [28, 39, 40].

Especially in the context of the dynamics of the production environment, this preceding step significantly increases epistemological complexity [1] which creates a much greater risk of making a wrong decision and ultimately taking an incorrect course of action. Polman suggests that this is even more the case if a decision is made for a third party [28]. This leads to the question of how to restrict the degrees of freedom in order to reduce the risk of making wrong decisions.

Basically, it can be said that too much complexity must always be avoided, otherwise the chain of action consisting of analysis and evaluation cannot be run through systematically [17, 29, 32, 33]. In addition, evaluating possible courses of action calls for an assessment of the consequences of the decision made. Therefore, only those degrees of freedom should be implemented whose benefits and consequences can be estimated by the decision-maker with a certain amount of confidence.

This leads to the following insight:

1. Additional degrees of freedom make a positive contribution to value creation if the type and number of resulting decision options do not disproportionately raise the risk of wrong decisions.
2. The question of the degrees of freedom required, or of the manageability by the decisive entity, always depends on the entity's competence.
3. The degree of overload can be strongly influenced by the way in which information is provided.

It has become apparent that, particularly for point 3, a large proportion of the fundamental logistical principles and main objectives can be transferred to handling information: the 5Rs of information logistics are therefore structured as follows: the right information (right choice), in the right quantity (right aggregation), at the right place, at the right time for the right decision-making entity [41, 42].

4. Organizational requirements for socio-technical systems

The requirements described above necessitate a socio-technical organizational structure that makes it possible to integrate people in terms of information and behavior so that they are cognitively addressed optimally and socially integrated. Only then they are able to make the right decisions at the right time and possess the necessary freedom and motivation to translate their decisions into correct action. This is confirmed by different studies in the field of cognitive ergonomics which state that mental workload can vary between low (i.e., underload) and very high levels (i.e., overload) [43, 44].

These two extremes are classified as inappropriate and can lead to imperfect or inaccurate perceptions, as well as to low levels of attention and capacity, and to insufficient time for a proper information processing [43–51]. High levels of mental workload occur when task demands exceed performer capacity [52]. [53]

Therefore, a location- and time-dependent resource allocation of data, knowledge, competencies, and performed actions is necessary. Additionally, requirements increase with the scope of the decision to be made. In the factory context, the smaller the quality control loop the better [54]. This also applies to decision-making control loops. Consequently, decisions affecting the direct creation of value should also be made close to the point where the value is added, since the information picture at this point is largely original and undistorted. This is underlined by the findings of Carley and Lin who state that regardless of information distortion performance is enhanced if there is a match between the complexity of organizational design an task environment [27].

In addition, the ability of the decision-making entity to recognize when it is overstretched, needs help, or can no longer assume responsibility should be considered. In the latter case, escalation management within a hierarchy chain is required. At the same time, the reverse case should also be taken into account. If the production-related person is capable of assuming responsibility for the right decision, he or she should be able and allowed to make it. This, in turn, requires the right of access to further information if the employee considers it useful for the decision. This increases the effectiveness and efficiency of value creation and is in line with the concept of minimizing the use of resources, since the information control loops should be kept as small as possible in the respective context. This also increases the employee's self-efficacy and enhances motivation and willingness to perform.

This context-related shifting of decision points along organizational structures is modelled in the MFF by means of flexible, heterarchical quality control loops.

In the MFF, information is provided in such a way that people can grasp the relevant facts quickly and correctly. It must therefore represent the facts correctly and, if possible, coherently. This particularly applies to the quality of the information. Contradicting information, or information that appears to contradict each other, leads to cognitive dissonance and leads to suboptimal decisions. Information must either be aligned, for example on different levels of aggregation, or must point these levels out. To avoid cognitive dissonance, it is important to ensure that information is not excessively distorted. The MFF embeds the information and the human being into the information supply by using the Standardized Coordinate System [20]. The manufacturing system is recorded by cameras and other sensors, merged with information from planning and control and the resulting information is then fed back into the real factory via projectors mounted on the hall ceiling (see Figure 3). This superposition of the digital and real world enables context-related information to be provided on the shopfloor in real time, thus supporting prompt decision-making adapted to the situation at hand. In addition, by overlapping the real factory with the digital image, cognitive dissonance is avoided because differing information is recognized immediately, enabling its cause to be ascertained. Furthermore, by interacting with the cameras on the hall ceiling, (e.g., via gestures), additional information can also be requested at any time, even when the system is in operation. The degree to which information is aggregated can be altered or the consequences of decisions can be simulated (e.g., when planning AGV routes).The MFF is increasingly using game engines which can quickly visualize even complex facts, simulate the consequences of actions and thus reduce the complexity of

Figure 3 Matrix Fusion Factory

decision-making [55–58]. The integration of information into the physical shopfloor means that effects can be viewed directly in the real system, thus enabling, for example, collisions to be predicted and avoided. In addition, the display of information via projectors also fits in with the concept of a minimum use of resources, since information and light are only coupled and projected where they are needed.

In the standardized coordinate system, all recorded data is classified according to time, place and position. It is then processed and, if it makes sense, merged or placed in relation to each other [20]. This also includes the information generated by the standardized coordinate system and projected into the shopfloor. The information provided by the projectors is simultaneously recorded again by the cameras in the context of, and together with, the real production. The MFF thus consists of a physical and a digital part of the manufacturing system, which together form the real manufacturing system on the shopfloor. This real-time fusion of real and digital facts, which also gives the MFF its name and enables correct action to be taken, can be used to identify distorted information or unmethodical actions before they lead to impairment of the entities or of value creation. This is true, whether the individual distorts the supplied information, the supplied information is distorted, or both. The state of the real manufacturing system can be retrieved in the standardized coordinate system for any point in time. Therefore, causes can be inferred by identifying the respective time and place. In an ideal case, the triggering action can be corrected in good time and the cause of the error eliminated before a negative effect can occur.

5. Technological requirements

The MFF strives to minimize the use of resources in value creation. To this end, only the number of resources deemed necessary to complete an order and optimize value creation are provided. Oversizing modular production resources is regarded as waste. For example, it makes no sense to use a six-axis robot if only three axes are required; the three remaining modules can be used to add value elsewhere.

In particular, the transformation enablers according to Nyhuis (scalability, modularity, mobility, compatibility and universality) [59] are prerequisites for the adaptability of a manufacturing system made up of groups of hardware and software modules. Classical production systems often consider transformation enablers primarily in a hardware context. A distinctive feature of the MFF is that the transformation enablers are also considered in the context of groups of software modules. Especially game engines have been able to demonstrate their capabilities in this regard [57].

However, transformation enablers are not sufficiently capable of determining how and to which extent an adjustment must be made to optimize value creation. This is in part due to the fact, that the effect of adjustments on the workload and capability of the individual decision-making entity is not considered.

To determine how the manufacturing system of ultra-flexible factories must be adapted to orders in a way that adds value, the information contained within the production system must be mapped in an up-to-date and accurate manner. Due to the system's dynamics and complexity, a map of this information can only be accurately generated if the system is able to describe itself in sufficient detail, at least in part. This presupposes the system's basic ability to describe itself.

For this the system must possess executable processes to describe itself and its parts which is especially important when new module groups with new process capabilities are assembled. The self-descriptive capability of the manufacturing system, including its process capabilities, is a prerequisite for achieving an optimal match between the order and the manufacturing system [1, 10, 11].

This technical self-description capability forms the basis of the socio-technical organizational structure. Since degrees of freedom always contribute effectively to value creation if the type and number of resulting decision options do not disproportionately increase the risk of wrong decisions, the complexity of the information map of the manufacturing system restricts the degrees of freedom that can be used to create value. This calls for a situation-specific representation of information depending on the deciding entity and the decision space available. The representation must be focused and sufficiently up to date to enable correct decisions to be made. These target variables conflict with each other, as shown in Figure 4 based on the target cross of production logistic variables.

The challenge in using computer-based information processing, also to support decisions made by humans, is the high degree of modeling effort required to represent complex systems. In recent years, the capabilities of machines and algorithms in terms of data processing, aggregation, and analysis have greatly improved. Especially the rapid development of artificial intelligence leads to a steady increase in performance [38, 60–62]. Nevertheless, complex simulation tasks often require costly or elaborate hardware equipment and extensive computation times, which conflicts with the requirements of near-real-time decision making. To enable the right decisions to be made at the right time, a balance between both objectives that is adapted to the situation must be achieved. Therefore, the accuracy of the model must be adjustable so that the decision maker is presented with a sufficiently complex picture of the facts within a sufficient period of time, thus allowing him or her to use the relevant information to make the right decision.

To this end, the MFF also uses computational decision-making processes such as machine-learning algorithms. These are mainly used to support human decision making by providing and preprocessing information, such as person-neutral detection of the viewing direction for optimal positioning and orientation of information projected onto the hall floor. Especially in dynamic and complex situations, the decision-making competence of humans is required. To avoid cognitive dissonance and improve the quality of decisions, in the MFF machine-processed information is integrated into the context of the real factory.

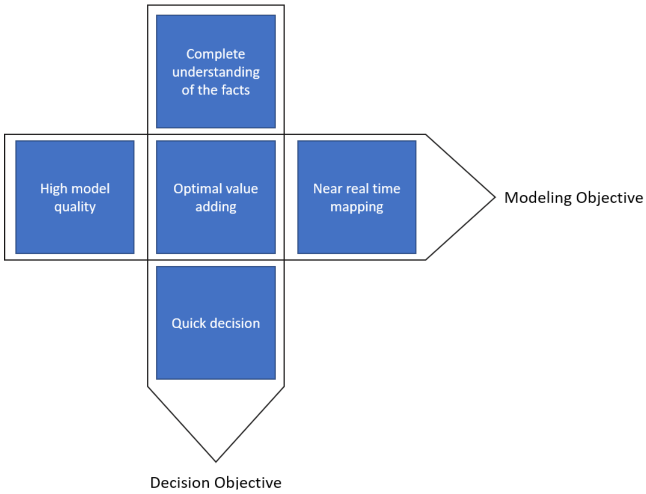


Figure 4 Target cross of information processing in production

This is in line with the position taken by the German Ethics Council, which points out the dubious quality of algorithm-based decision support systems and is concerned about effects on human self-efficacy [63]. In the MFF, therefore, the focus is on the ability of the person to adapt the representation of information to his or her needs and to understand and verify it. This improves the receptivity and quality of decisions and enables the person to be more involved in decision-making processes. The standardized coordinate system also makes it possible to document and assign decisions.

6. Conclusion and Future Works

The paper has demonstrated the role of information and information distortion in ultra-flexible factories, particularly with regard to the concept of minimizing the use of resources. It has also explained how permissible information distortion can be determined. Based on this, the role of human and machine decision making was considered, the requirements of complex information management were ascertained and presented, and the effects on the organizational structure and technology were derived. The target cross of production information was developed and introduced.

In the future, approaches will be investigated that improve model quality and reaction times in the context of the provision of information by machines. In addition, further cameras and sensors on different levels of the MFF will be included in the Standardized Coordinate System. The integration of additional data processing capabilities into a holistic information logistics approach will be examined. As part of this, the

5R approach is to be validated. Also, greater emphasis will be placed on the interaction between humans and technology. Exoskeletons also will be integrated into the MFF and the concepts presented in this paper will be transferred to human-exoskeleton interaction.

References

[1] J. Siegert, T. Schlegel, L. Zarco, B. Miljanovic, A. Meyke, and T. Bauernhansl, "Ultra-flexible Factories: An Approach to Manage Complexity," Procedia CIRP, vol. 93, pp. 329–334, 2020, doi: 10.1016/j.procir.2020.04.112.

[2] J. Siegert, T. Schlegel, and T. Bauernhansl, "Matrix Fusion Factory," Procedia Manufacturing, vol. 23, pp. 177–182, 2018, doi: 10.1016/j.promfg.2018.04.013.

[3] A. Kluth, J. Jäger, A. Schatz, and T. Bauernhansl, "Method for a Systematic Evaluation of Advanced Complexity Management Maturity," Procedia CIRP, vol. 19, pp. 69–74, 2014, doi: 10.1016/j.procir.2014.05.041.

[4] Andreas Kluth, Jens Jäger, Anja Schatz, and Thomas Bauernhansl, "Evaluation of Complexity Management Systems – Systematical and Maturity-based Approach," Procedia CIRP, vol. 17, pp. 224–229, 2014, doi: 10.1016/j.procir.2014.01.083.

[5] H. ElMaraghy et al., "Product variety management," CIRP Annals, vol. 62, no. 2, pp. 629–652, 2013, doi: 10.1016/j.cirp.2013.05.007.

[6] S. Mattsson, E. Ekstrand, and M. Tarrar, "Understanding disturbance handling in complex assembly: analysis of complexity index method results," Procedia Manufacturing, vol. 25, pp. 213–222, 2018, doi: 10.1016/j.promfg.2018.06.076.

[7] B. Brinzer and K. Schneider, "Complexity assessment in production: linking complexity drivers and effects," Procedia CIRP, vol. 93, pp. 694–699, 2020, doi: 10.1016/j.procir.2020.04.014.

[8] V. Modrak, D. Marton, and S. Bednar, "The Influence of Mass Customization Strategy on Configuration Complexity of Assembly Systems," Procedia CIRP, vol. 33, pp. 538–543, 2015, doi: 10.1016/j.procir.2015.06.071.

[9] B. Alkan, D. A. Vera, M. Ahmad, B. Ahmad, and R. Harrison, "Complexity in manufacturing systems and its measures: a literature review," EJIE, vol. 12, no. 1, p. 116, 2018, doi: 10.1504/EJIE.2018.089883.

[10] T. Schlegel, J. Siegert, and T. Bauernhansl, "Metrological Production Control for Ultra-flexible Factories," Procedia CIRP, vol. 81, pp. 1313–1318, 2019, doi: 10.1016/j.procir.2019.04.019.

[11] T. Schlegel, J. Siegert, T. Mahr, L. Zarco, F. Herbrig, and T. Bauernhansl, "Metrological PPC: Determination of added-value driven module process graphs," Procedia CIRP, vol. 93, pp. 634–639, 2020, doi: 10.1016/j.procir.2020.04.123.

[12] J. Siegert, T. Schlegel, L. Zarco, and T. Bauernhansl, "Order-Oriented Learning Factories: Why and How Learning Factories Have to Adapt," Procedia Manufacturing, vol. 45, pp. 460–465, 2020, doi: 10.1016/j.promfg.2020.04.053.

[13] T. Bauernhansl et al., "Semantic Structuring of Elements and Capabilities in Ultra-flexible Factories," Procedia CIRP, vol. 93, pp. 335–340, 2020, doi: 10.1016/j.procir.2020.04.010.

[14] P. Milling, Systemtheoretische Grundlagen zur Planung der Unternehmenspolitik. Zugl.: Mannheim, Univ., Habil.-Schr. Berlin: Duncker & Humblot, 1981.

[15] W. Weaver, "Science and complexity," American scientist, vol. 36, no. 4, pp. 536–544, 1948.

[16] V. Ferretti, Back to Ptolemaism: To Protect the Human Individual from Abuses of Social Constructs: Print-on-Demand, 2012.

[17] J. Siegert, T. Schlegel, and T. Bauernhansl, “Verifiable Competencies for Production Technology,” Procedia Manufacturing, vol. 45, pp. 466–472, 2020, doi: 10.1016/j.promfg.2020.04.054.

[18] P. Mandal and K. Bagchi, “Strategic role of information, knowledge and technology in manufacturing industry performance,” Industrial Management & Data Systems, vol. 116, no. 6, pp. 1259–1278, 2016, doi: 10.1108/IMDS-07-2015-0297.

[19] T. Dewett and G. R. Jones, “The role of information technology in the organization: a review, model, and assessment,” Journal of Management, vol. 27, no. 3, pp. 313–346, 2001, doi: 10.1016/S0149-2063(01)00094-0.

[20] J. Siegert, T. Schlegel, E. Groß, and T. Bauernhansl, “Standardized Coordinate System for Factory and Production Planning,” Procedia Manufacturing, vol. 9, pp. 127–134, 2017, doi: 10.1016/j.promfg.2017.04.032.

[21] M. Schreiber, P. Bausch, J. Best, and J. Metternich, “Datenanalyse in Produktionsprozessen,” ZWF, vol. 115, no. 5, pp. 309–313, 2020, doi: 10.3139/104.112285.

[22] K. Rönick, B. Röhm, P. Bausch, R. Anderl, and J. Metternich, “Digitale Informationsbereitstellung in der Produktion,” ZWF, vol. 115, no. 6, pp. 438–441, 2020, doi: 10.3139/104.112372.

[23] A.-S. Chaxel, C. Wiggins, and J. Xie, “The impact of a limited time perspective on information distortion,” Organizational Behavior and Human Decision Processes, vol. 149, pp. 35–46, 2018, doi: 10.1016/j.obhdp.2018.09.001.

[24] E. Polman, “Information distortion in self-other decision making,” Journal of Experimental Social Psychology, vol. 46, no. 2, pp. 432–435, 2010, doi: 10.1016/j.jesp.2009.11.003.

[25] J. Russo, V. H. Medvec, and M. G. Meloy, “The Distortion of Information during Decisions,” Organizational Behavior and Human Decision Processes, vol. 66, no. 1, pp. 102–110, 1996, doi: 10.1006/obhd.1996.0041.

[26] S. D. Bond, K. A. Carlson, M. G. Meloy, J. E. Russo, and R. J. Tanner, “Information distortion in the evaluation of a single option,” Organizational Behavior and Human Decision Processes, vol. 102, no. 2, pp. 240–254, 2007, doi: 10.1016/j.obhdp.2006.04.009.

[27] K. M. Carley and Z. Lin, “A Theoratical Study of Organizational Performance Under Information Distortion,” MANAGEMENT SCIENCE, vol. 1997, Vol. 43, No.7, pp. 976–997, 1997.

[28] E. Polman, “Effects of self-other decision making on regulatory focus and choice overload,” Journal of personality and social psychology, vol. 102, no. 5, pp. 980–993, 2012, doi: 10.1037/a0026966.

[29] M. R. Endsley, “Toward a Theory of Situation Awareness in Dynamic Systems,” Hum Factors, vol. 37, no. 1, pp. 32–64, 1995, doi: 10.1518/001872095779049543.

[30] F. A. Csaszar and J. P. Eggers, “Organizational Decision Making: An Information Aggregation View,” MANAGEMENT SCIENCE, vol. 59, no. 10, pp. 2257–2277, 2013, doi: 10.1287/mnsc.1120.1698.

[31] L. Bouargane and A. Cherkaoui, “Towards an explicative model of human cognitive process in a hidden hazardous situation and a cognitive ergonomics intervention in railway environment,” in 2015 International Conference on Industrial Engineering and Systems Management (IESM), Seville, Spain, Oct. 2015 - Oct. 2015, pp. 968–976.

[32] M. R. Endsley, “Situation awareness in aviation systems,” in Human factors in transportation, Handbook of aviation human factors, Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 1999, pp. 257–276.

[33] M. Endsley, “Situation awareness analysis and measurement, chapter theoretical underpinnings of situation awareness,” A Critical Review, pp. 3–33, 2000.

[34] Vidulich, M., Dominguez, C., Vogel, E., McMillan, G., Situation Awareness: Papers and Annotated Bibliography., 1994.

[35] M. Hoedemaeker, Summary Description of Workload Indicators: WP1 Workload Measures. Human Machine Interface and the Safety of Traffic in Europe Growth Project: HASTE. Leeds UK: GRD1-2000-25361, 2002.

[36] J. Hiba, H. Hadi, A. Hameed Shnain, S. Hadishaheed, and A. Haji, “BIG DATA AND FIVE V’S CHARACTERISTICS,” pp. 2393–2835, 2015.

[37] J.-C. Pomerol, “Artificial intelligence and human decision making,” European Journal of Operational Research, vol. 99, no. 1, pp. 3–25, 1997, doi: 10.1016/S0377-2217(96)00378-5.

[38] F. Emmert-Streib, O. Yli-Harja, and M. Dehmer, “Artificial Intelligence: A Clarification of Misconceptions, Myths and Desired Status,” Frontiers in artificial intelligence, vol. 3, p. 524339, 2020, doi: 10.3389/frai.2020.524339.

[39] A. Falck, Roland Örtengren, and M. Rosenqvist, “Relationship between complexity in manual assembly work, ergonomics and assembly quality,” in 2012.

[40] Y. Wang and G. Ruhe, “The Cognitive Process of Decision Making,” International Journal of Cognitive Informatics and Natural Intelligence, vol. 1, no. 2, pp. 73–85, 2007, doi: 10.4018/jcini.2007040105.

[41] D. M. Haftor, M. Kajtazi, and A. Mirijamdotter, “A Review of Information Logistics Research Publications,” in Business Information Systems Workshops, 2011, pp. 244–255.

[42] Barbara Dinter, “Success factors for information logistics strategy — An empirical investigation,” Decision Support Systems, vol. 54, no. 3, pp. 1207–1218, 2013, doi: 10.1016/j.dss.2012.09.001.

[43] K. A. Brookhuis and D. Waard, “Monitoring drivers' mental workload in driving simulators using physiological measures,” Accident; analysis and prevention, vol. 42, no. 3, pp. 898–903, 2010, doi: 10.1016/j.aap.2009.06.001.

[44] De Waard, “On the measurement of driver mental workload,” p. 161, 1997.

[45] M. G. Lenné, T. J. Triggs, and J. R. Redman, “Time of day variations in driving performance,” Accident Analysis & Prevention, vol. 29, no. 4, pp. 431–437, 1997, doi: 10.1016/S0001-4575(97)00022-5.

[46] S. Leung and G. Starmer, “Gap acceptance and risk-taking by young and mature drivers, both sober and alcohol-intoxicated, in a simulated driving task,” Accident Analysis & Prevention, vol. 37, no. 6, pp. 1056–1065, 2005, doi: 10.1016/j.aap.2005.06.004.

[47] L. N. Boyle, J. Tippin, A. Paul, and M. Rizzo, “Driver Performance in the Moments Surrounding a Microsleep,” Transportation Research Part F: Traffic Psychology and Behaviour, vol. 11, no. 2, pp. 126–136, 2008, doi: 10.1016/j.trf.2007.08.001.

[48] T. Nilsson, T. M. Nelson, and D. Carlson, “Development of fatigue symptoms during simulated driving,” Accident Analysis & Prevention, vol. 29, no. 4, pp. 479–488, 1997, doi: 10.1016/S0001-4575(97)00027-4.

[49] M. E. Rakauskas, N. J. Ward, E. R. Boer, E. M. Bernat, M. Cadwallader, and C. J. Patrick, “Combined effects of alcohol and distraction on driving performance,” Accident; analysis and prevention, vol. 40, no. 5, pp. 1742–1749, 2008, doi: 10.1016/j.aap.2008.06.009.

[50] P. Thiffault and J. Bergeron, “Monotony of road environment and driver fatigue: a simulator study,” Accident Analysis & Prevention, vol. 35, no. 3, pp. 381–391, 2003, doi: 10.1016/S0001-4575(02)00014-3.

[51] W. B. Verwey and D. M. Zaidel, “Preventing drowsiness accidents by an alertness maintenance device,” Accident Analysis & Prevention, vol. 31, no. 3, pp. 199–211, 1999, doi: 10.1016/S0001-4575(98)00062-1.

[52] S. Loft, P. Sanderson, A. Neal, and M. Mooij, “Modeling and predicting mental workload in en route air traffic control: critical review and broader implications,” Hum Factors, vol. 49, no. 3, pp. 376–399, 2007, doi: 10.1518/001872007X197017.

[53] F. P. Da Silva, “Mental Workload, Task Demand and Driving Performance: What Relation?,” Procedia - Social and Behavioral Sciences, vol. 162, pp. 310–319, 2014, doi: 10.1016/j.sbspro.2014.12.212.

[54] M. Kandler, J. J. Kurtz, N. Stricker, and G. Lanza, “Digitales, agiles Management auf dem Shopfloor,” ZWF, vol. 115, 1-2, pp. 23–26, 2020, doi: 10.3139/104.112241.

[55] M. Paul, D. Leiber, J. Pleli, and G. Reinhart, Simulation of Reconfigurable Assembly Cells with Unity3D, IFIP Advances in Information and Communication Technology: Springer International Publishing, vol. 566.

[56] J. Siegert, L. Zarco, T. Rossmeissl, and T. Schlegel, “Software Control for a Cyber-Physical System in a Manufacturing Environment based on a Game Engine,” in 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, Jun. 2019 - Jun. 2019, pp. 1–7.

[57] J. Siegert, L. Zarco, T. Schlegel, and T. Bauernhansl, "Software Control System Requirements for Ultra-Flexible Learning Factories," Procedia Manufacturing, vol. 45, pp. 442–447, 2020, doi: 10.1016/j.promfg.2020.04.050.

[58] J. Siegert, L. Zarco, and T. Schlegel, "Universal Accessibility Concept for Controlling Production Means in Manufacturing Systems," in 2020 IEEE International Conference on Industrial Technology (ICIT), Buenos Aires, Argentina, Feb. 2020 - Feb. 2020, pp. 349–354.

[59] P. Nyhuis, T. Heinen, G. Reinhart, C. Rimpau, E. Abele, and A. Wörn, "Wandlungsfähige Produktionssysteme. Theoretischer Hintergrund zur Wandlungsfähigkeit von Produktionssystemen," Werkstattstechnik online, 98 (1-2), pp. 85–91, 2008.

[60] J. F. Arinez, Q. Chang, R. X. Gao, C. Xu, and J. Zhang, "Artificial Intelligence in Advanced Manufacturing: Current Status and Future Outlook," Journal of Manufacturing Science and Engineering, vol. 142, no. 11, 2020, doi: 10.1115/1.4047855.

[61] C.-F. Chien, S. Dauzère-Pérès, W. T. Huh, Y. J. Jang, and J. R. Morrison, "Artificial intelligence in manufacturing and logistics systems: algorithms, applications, and case studies," International Journal of Production Research, vol. 58, no. 9, pp. 2730–2731, 2020, doi: 10.1080/00207543.2020.1752488.

[62] A. Kusiak, "Convolutional and generative adversarial neural networks in manufacturing," International Journal of Production Research, vol. 58, no. 5, pp. 1594–1604, 2020, doi: 10.1080/00207543.2019.1662133.

[63] Deutscher Ethikrat, Wie wird die Zukunft mit intelligenten Maschinen? Berlin, 2017. Accessed: Dec. 12 2020. [Online]. Available: https://www.ethikrat.org/mitteilungen/2017/wie-wird-die-zukunft-mit-intelligenten-maschinen/

Biography

Jörg Siegert (*1978) has been the Senior Engineer and Deputy Head of the Institute of Industrial Manufacturing and Management IFF since 2012: He has also been acting deputy head of the newly founded Institute of Energy Efficiency in Production EEP since 2012.

Thilo Schlegel (*1985) graduated in industrial engineering at the University of Stuttgart and is working as a research assistant for Fraunhofer Institute for Manufacturing Engineering and Automation IPA and the Institute of Industrial Manufacturing and Management IFF where he heads the Matrix Fusion Factory since 2019.

Liliana Zarco (*1989) graduated in bionic engineering at the National Polytechnic Institute of Mexico and master of industrial business and management at the University of Applied Sciences Esslingen. Since 2018, she is doctoral candidate and research assistant for the Institute of Industrial Manufacturing and Management IFF on the topics of ontology, gamification and information models of the adaptive factory.

Thomas Bauernhansl (*1969) has been the director of the Institute of Industrial Manufacturing and Management IFF at the University of Stuttgart and the Fraunhofer Institute for Manufacturing Engineering and Automation IPA since 2011. He is a member of the WGP, the Scientific Society for Production Technology as well as Deputy Chairman of the Steering Committee of Allianz Industrie 4.0 BW.

2nd Conference on Production Systems and Logistics

Classification Model Of Supply Chain Events Regarding Their Transferability To Blockchain Technology

Fabian Dietrich[1,2], Eldin Mehmedovic[1,2], Ali Turgut[3], Louis Louw[2], Daniel Palm[1,4]

[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]*Steinbeis Innovation gGmbH, Alteburgstr. 150, 72762 Reutlingen, Germany*
[4]*Fraunhofer Institute for Manufacturing Engineering and Automation, Alteburgstr. 150, 72762 Reutlingen, Germany*

Abstract

The blockchain technology represents a decentralized database that stores information securely in immutable data blocks. Regarding supply chain management, these characteristics offer potentials in increasing supply chain transparency, visibility, automation, and efficiency. In this context, first token-based mapping approaches exist to transfer certain supply chain events to the blockchain, such as the creation or assembly of parts as well as their transfer of ownership. However, the decentralized and immutable structure of blockchain technology also creates challenges. In particular, the scalability, storage capacity, and the special requirements for storage formats make it currently impossible to map all supply chain events unrestrictedly on the blockchain. As a first step, this paper identifies important supply chain events for different use cases combining blockchain technology and supply chain management. Secondly, the supply chain events are classified in terms of their expected technical properties and their relevance for the respective use case. Finally, the identified supply chain events are evaluated regarding their transferability to blockchain technology and a classification model is introduced.

Keywords

Blockchain; Supply Chain Events; Supply Chain Management; Classification Model

1. Introduction

To overcome the challenges in supply chains, such as increasing transparency, traceability, security, etc., the potential of blockchain technology has been proven in recent years [1]. The blockchain technology represents a distributed electronic register to store and retrieve information permanently, transparently, and in a trustworthy manner without the need of accessing a central instance [2,3]. Information is stored in so-called 'blocks'. Thereby the blocks and thus also the information in the blocks are encrypted with so-called 'hashes' [4]. Hashes are an encryption and can be seen as a kind of fingerprint or identification of the individual block. When a new block is created in this chain, a new hash is generated for the emerging block. In addition to its own hash, a block always includes the hash of the previous block. Therefore, the individual blocks are connected in a linear sequence to form a secure chain [4]. Another significant feature of blockchain technology is the decentralised network whereby every network participant has access to the complete blockchain [3]. When a new block is created, this information is sent to all participants in the network, so each of these participants verifies the newly created block and confirms the conducted changes in this block [5]. The blockchain technology is thus a decentralised database that can hardly be manipulated and is accessible to every participant in the network [4]. With the implementation of smart contracts, the technology became relevant for application areas beyond the financial sector [6]. Smart contracts are prefabricated program codes that specify the contract terms of the respective contract partners in advance and are triggered when the contract terms are fulfilled [6]. The mentioned characteristics of the blockchain in combination with smart contracts can be used for example to overcome challenges of supply chains [7,8] to increase their transparency, traceability, automation, and immutability [1]. A major challenge, however, is the simultaneous use of the three properties: decentralisation, security, and scalability. For example, increasing the degree of decentralisation and security of a blockchain system automatically reduces its scalability. This current correlation is also known as the blockchain trilemma [9]. In order to circumvent the scalability problem, it requires to divide supply chain scenarios into supply chain events (SCEs) to differentiate them individually when transferring them to a system such as blockchain technology [10]. The "Event Product Code Information Service (EPCIS) enables different applications to create and share event data with visibility, both within an enterprise and across enterprises [11]". Each event is described with the four dimensions of the generic 'EPCIS-Event'. These four dimensions are the object(s) or other entity(ies) that are the subject of the event, the date and time, the location where the event occurred and the business context [11]. In summary, the generic EPCIS-Event describes "what, when, where and why" happens. The 'what', in turn, is differentiated into the four core event types Object Event, Aggregation Event, Transaction Event and Transformation Event [11]. According to GS1, the definitions of these core SCEs are as follows [11]:

- *Object Event*. "The Object Event represents an event that has happened to one or more physical or digital objects".
- *Aggregation Event*. "The Aggregation Event represents an event that has happened to one or more objects that are physically aggregated (physically forced to be in the same place at the same time)".
- *Transaction Event*. "The Transaction Event represents an event where one or more objects are associated or disassociated with one or more identified business transactions".
- *Transformation Event*. "The Transformation Event represents an event in which input objects are fully or partially consumed and output objects are produced, such that any of the input objects may have contributed to all of the output objects".

This paper aims to classify SCEs regarding their transferability in a blockchain technology context. First, challenges when transferring SCEs to blockchain technology are presented to raise problem awareness and explain the rationale of the research. Secondly, the classification model is developed based on the fundamental structure of blockchain architectures in supply chain management and relevant data quality dimensions. Finally, a classification model is proposed and the results are discussed.

2. Related works and rationale of the research

Especially when mapping supply chains with complex products on the blockchain, a transfer of different SCEs is necessary [1,12]. To enable a holistic mapping, GS1 defines generic EPCIS-Events consisting of Object Events, Aggregation Events, Transaction Events, and Transformation Events [11]. With its special decentralised and immutable structure, the blockchain technology poses special requirements for the transfer of SCEs, which have not yet been uniformly clarified in research. Holtkemper [12] considers the storing of SCEs with (large) files on the blockchain practical at certain stages in the supply chain. Even though Holtkemper [12] considers the storage of files generally to be possible, he points out that due to the increasing storage capacity of the blockchain transferring certain files to different storage systems could be reasonable. Hepp et al. [13], on the other hand, point out that blockchain's structure has such a limiting effect on the transfer of SCE-Data, that certain files and transactions cannot be stored on the blockchain. In addition, Hepp et al. [13] indicate that the localisation of the data to be stored has an impact on the data integrity and transparency of the overall system. To clarify this matter, this paper develops a classification model of SCEs regarding their transferability to blockchain technology. In order to develop the model, data quality dimensions of SCEs are identified and evaluated in a blockchain technology context. The proposed classification model is intended to serve as a foundation for evaluating supply chain use cases in terms of their sense and feasibility with the adoption of blockchain technology.

3. Development of a classification model

In this chapter the development of the classification model takes place. First, the fundamental characteristics of blockchain architectures in supply chain management are described. Subsequently, this chapter discussed the data quality dimensions to be considered when transferring SCEs to blockchain technology. Finally, the results are summarised in a classification model.

3.1 Blockchain architectures in supply chain management

A holistic supply chain mapping has to include the mapping of the four supply chain core events (Object, Aggregation, Transformation and Transaction Events) [11]. As the evaluation of blockchain publications in supply chain management shows, most of the research deals with mapping approaches based on simple supply chains [1]. This includes the mapping of subsets of the existing event categories mentioned above, but not a holistic mapping of all core events in one architecture [1]. For this to be possible, it requires to connect physical assets to unique identifiers or 'digital profiles' on the blockchain [14], which is also known as 'tokenising of assets' [15]. Fundamentally, one can distinguish between two types of tokens, fungible token and non-fungible token [22]:

- Fungible token (FT): Different units of a FT are interchangeable and have no unique properties.
- Non-fungible token (NFT): Each unit of a NFT is unique from another, allowing the tracking of their ownership.

Based on these technical properties, FTs are particularly suitable to represent cryptocurrencies while NFTs, as the term 'non-fungibility' suggests, are intended to identify unique digital or non-digital assets [16]. This makes the adoption of NFTs suitable for mapping assets throughout supply chains on the blockchain [17]. For the holistic mapping of complex supply chains two different token-based approaches exist.

The first approach aims to map SCEs in individual smart contracts based on the non-fungible ERC-721 token standard [18]. Each SCE, except Transaction Events, is represented by a smart contract on the blockchain network. Transaction Events are already embedded into the ERC-721 token standard and do not require the deployment of own 'transaction' smart contracts. The other SCEs are represented by individual smart contracts and their creation can be connected to certain conditions. This means that the creation of a token can be dependent on another token, but the tokens themselves, however, are not logically coupled [19]. This architecture with individual smart contracts without logical coupling enables a fast implementation but merely a static mapping of supply chains on the blockchain [20]. Consequently, dynamic adjustments in the supply chain or product structure lead to a redeployment of all smart contracts which are affected directly or indirectly (smart contracts with defined dependencies on each other). This makes it difficult to maintain the architecture in supply chains that place high demands in terms of continuous flexible adjustments [20]. In addition, the lack of logical coupling complicates considerably the auditability of the individual tokens [19].

The second approach aims to map all SCEs holistically with functions embedded into one smart contract [20]. The individual SCEs can be both, dependent on each other through defined conditions and logically coupled with each other on an application level. With the help of an embedded authority concept, dynamic adjustments in the supply chain or product structure can be conducted without the need of redeploying a new smart contract on the blockchain network [20]. This allows the mapping of supply chains that place high demands in terms of continuous flexible adjustments. However, this approach considerably increases the planning effort before the deployment of the smart contract. Due to the immutability of the blockchain, eventualities regarding the structure and the supply chain events must be defined in advance and embedded into the source code [20]. Figure 1 shows the fundamental architecture with one supply chain smart contract in a decentralised blockchain network.

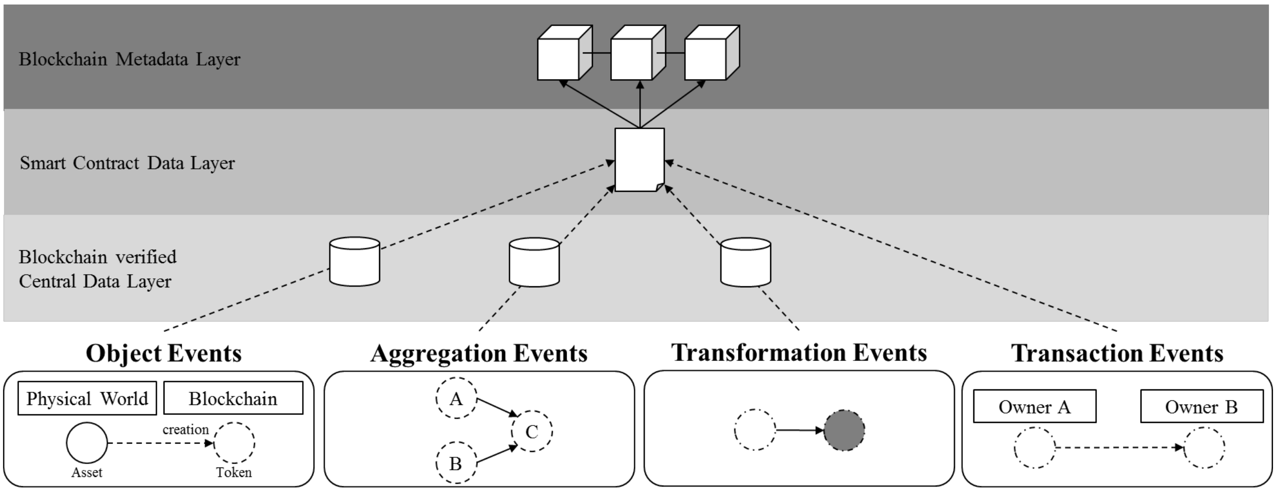


Figure 1: Fundamental architecture with one supply chain smart contract in a decentralised blockchain network

Figure 1 shows that in such architecture one can distinguish between three different data layers. The blockchain metadata layer provides the basis. Depending on the type of blockchain, private or public, either only selected participants have access to this data layer, or it is publicly accessible. The blockchain metadata layer includes the metadata of the network such as the time of transactions, the transaction initiators, or the type of transaction (financial transaction or smart contract interaction) [21]. The second layer consists of the smart contract data layer. This data layer contains the data to be stored within the smart contract. Only authorities with access to the smart contract have the possibility to access this information or add new information. From the perspective of transferring SCEs to the blockchain, this data layer represents the critical data storage. Unlike the metadata layer, the smart contract developer can determine the scope of data to be stored in this layer. However, they must adhere to the framework conditions of the blockchain metadata layer (see Section 3.2). Due to the limited scalability of blockchain technology, various approaches exist that extend the smart contract data layer with a blockchain-verified central data layer – so-called blockchain verified off-chain storages [21,22]. Such off-chain storage approaches store data in a content-addressable storage system, which stores files by their hashes. The respective hash is also stored in the smart contract on the blockchain and serves as a reference to the complete data set. Clients can then compare the hash resulting from the off-chain data with the hash stored in the smart contract on-chain. A modification of the data would result in another hash making it easy to verify the correctness of the off-chain stored data [22].

This raises the critical question of what data should or can be stored on-chain in the smart contract data layer and what data should be transferred to the blockchain verified central data layer. However, the type of SCE itself cannot determine this. For example, for some Object Events it is sufficient to add a simple identification number to the object when creating it; for other Object Events it is necessary to add an entire construction plan or a CAD drawing [12]. As a basis for such decisions, this paper considers an evaluation of SCEs based on different data quality dimensions.

3.2 Data quality dimensions

The mapping of core SCEs represents a transfer of data from multiple sources in a larger information systems context. When conceptualizing such larger information systems, Strong et al. [23] recommend a holistic view on data qualities (DQs). In this context, larger information systems “cover the organizational processes, procedures, and roles employed in collecting, processing, distributing and using data [23]“.

Strong et al. [23] define four DQ categories: Accessibility, Representational, Contextual, and Intrinsic. The Accessibility DQ and the Representational DQ are technical in nature and must be ensured within the information system. The Contextual DQ and the Intrinsic DQ, however, must be ensured during the transfer of data into the information system. Transferred to the blockchain architecture presented in Section 3.1, the Accessibility DQ and the Representational DQ must be ensured within the blockchain technology and the smart contract, while the Contextual DQ and the Intrinsic DQ must be ensured when transferring SCE-Data to the blockchain and the smart contract.

Table 1: DQ categories and dimensions with blockchain relevance

DQ Category	DQ Dimensions
Contextual DQ	Relevancy, Value-Added, Timelessness, Completeness, Amount of data
Intrinsic DQ	Accuracy, Objectivity, Believability, Reputation
Blockchain Suitability DQ	Data Format, Transaction Frequency, Data Size

This paper extends the DQ categories by Strong et al. [23] with a blockchain specific DQ which aims to evaluate the suitability for storing data on-chain. Table 1 shows the DQ categories and dimensions with blockchain relevance. In this context, important DQ dimensions are represented by the data format (transactions, source code, or files) [12], the transaction frequency as well as the data size [13].

- *Data format*. Current blockchain platforms support the storage of transactions and codes, but do not support the storage of files [12]. Approaches exist that convert files into a blockchain transaction format, store them on-chain, and decode them to restore the file at any point [24]. However, if it is necessary that certain files remain in a certain format (e.g. PDF-files) throughout the supply chain, it is not suitable to store these files on-chain and a blockchain-verified off-chain storage is preferable.
- *Transaction frequency*. Blockchain platforms are inferior to central systems in terms of their transaction processing. For example, the main Ethereum network currently processes only 14 transactions per second [25]. Therefore, SCEs that require a higher scalability regarding the transaction processing, cannot be transferred to the blockchain network. If the transaction frequency of a single SCE is just below the transactions-per-second limit, it can theoretically be supported by the network, but could have a negative impact on the usability of the overall system. In this case, the overall transaction load of all SCEs must be aligned with the network capacity and considered in the decision whether events should be stored on-chain or off-chain. In this context, the threshold values differ depending on the blockchain platform used.
- *Data size*. The data size that can be transferred to the blockchain within one transaction is limited by the block size of the used blockchain platform [13]. Currently, the average block size of the Ethereum blockchain is around 50 KB. Similar to the transaction frequency, transactions with bigger data size must be evaluated regarding the holistic SCE context to determine whether the transfer of such events could have a negative impact on the usability of the overall system. Again, the threshold values also differ depending on the blockchain platform used.

The findings regarding the Blockchain Suitability DQ together with the DQs listed in Table 2 are transferred to a 3-level blockchain-related model (see Figure 2). The Contextual DQ ranges from irrelevant to relevant for the respectively considered use case. In between are SCE-Data, which are not relevant for the actual use case but potentially offer added value to the considered use case (e.g. detailed shipping information). The Intrinsic DQ ranges from poor quality to high quality, whereby the lowest DQ is for example represented by manual data input and the highest DQ can be provided by so-called blockchain oracles. “A blockchain oracle is a mechanism that fetches data from the external world to include it in the isolated execution environment of a blockchain. [...] Blockchain oracles are needed to bridge blockchains and the external world because of unique characteristics of blockchain [26]”. Therefore, blockchain oracles act as trusted intermediary service between blockchains and a variety of independent data sources [27]. In between, ‘normal’ IoT (Internet-of-Things) devices that can transfer their data automatically to the blockchain provide a medium Intrinsic DQ. Figure 2 shows a summary of the findings with a three-level subdivision of the DQ categories. The results of this chapter are summarised and discussed in a classification model in Section 3.3.

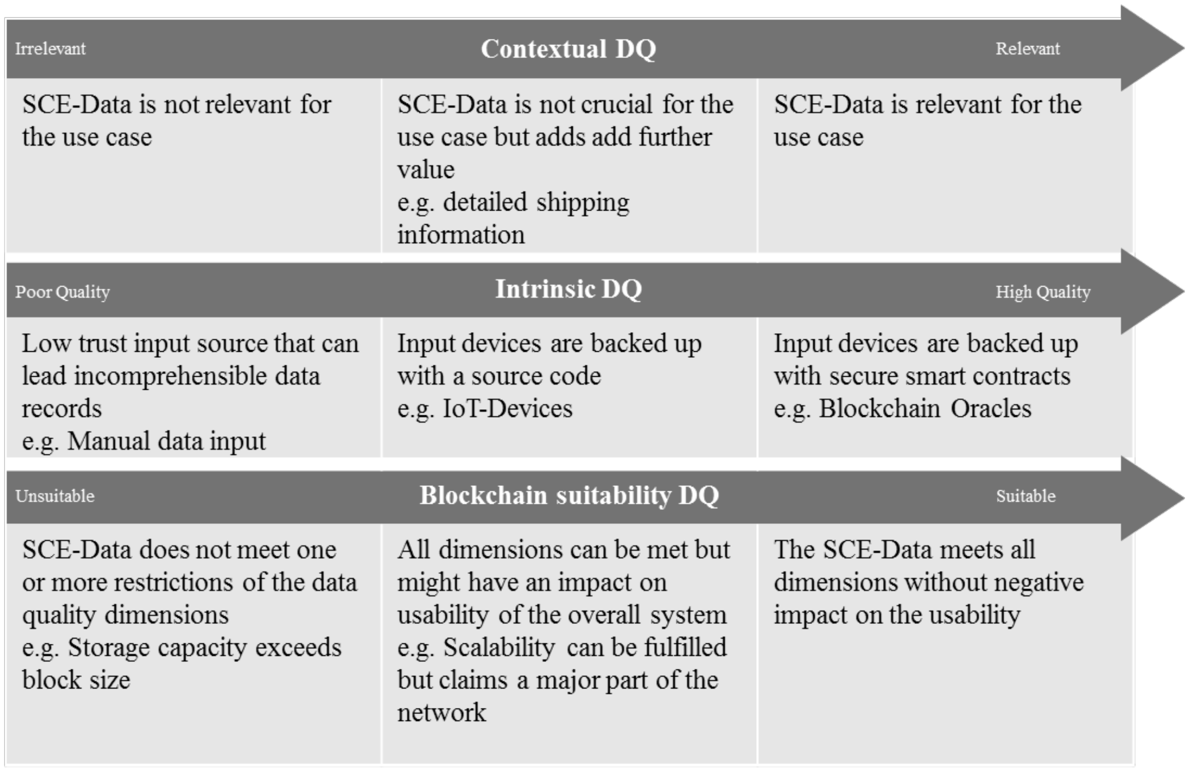


Figure 2: Three-level subdivision of the DQ categories with blockchain relevance

3.3 Classification Model

The findings of Chapter 3 form the basis for the three-dimensional classification model of supply chain events regarding their transferability to blockchain technology. Figure 3 shows the proposed classification model. In this model, the z-axis represents the Contextual DQ, the x-axis the Intrinsic DQ, and the y-axis the Blockchain (BC) Suitability DQ. The first number (1 x x) indicates the Contextual DQ, whereby irrelevant SCEs are rated with a ‘3’ and relevant SCEs are rated with a ‘1’. These SCEs have add either low or high value to the respective use case. The second number (x 1 x) indicates the Intrinsic DQ, whereby poor quality SCEs are rated with a ‘3’ and high quality SCEs are rated with a ‘1’. These SCEs determine the integrity of the input data and therefore affect the integrity of the overall system.

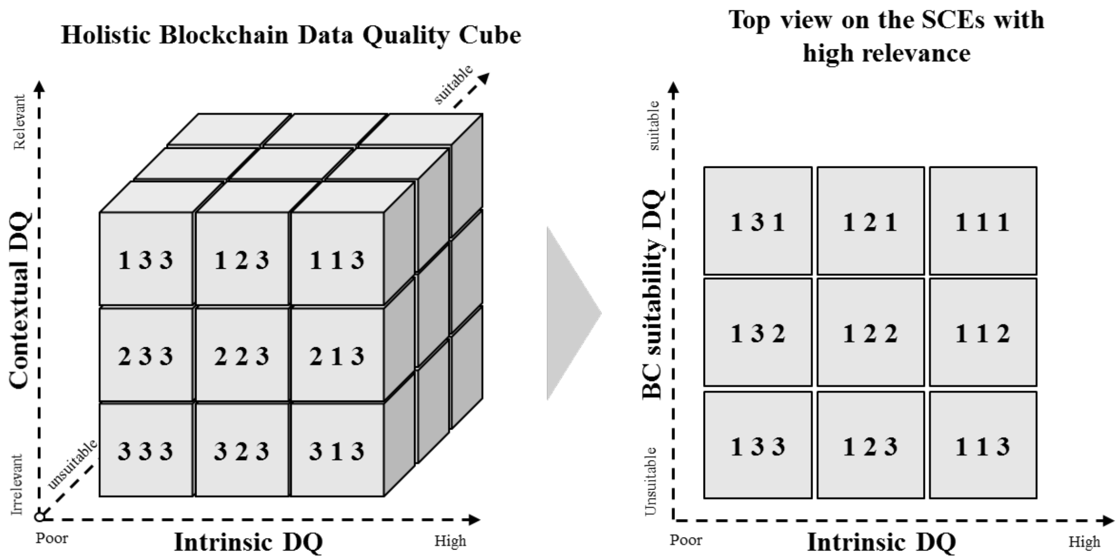


Figure 3: Classification model of SCEs regarding their transferability to blockchain technology

The third number (x x 1) indicates the BC Suitability DQ, whereby unsuitable SCEs are rated with a '3' and suitable SCEs are rated with a '1'. These SCEs determine the data transparency of the system. SCEs with are unsuitable for on-chain storage must be stored in blockchain verified off-chain storages and therefore indicate a low data transparency. In reverse, SCEs that are suitable for an on-chain storage increase the data transparency of the system. As the model shows, the SCEs events with low use case relevance, poor Intrinsic DQ, and unsuitable for an on-chain storage (3 3 3) represent the worst classification. The counterpart is represented by SCEs with high use case relevance, high Intrinsic DQ, and high suitability for storing them on-chain (1 1 1). Such SCEs add high value to the respective use case, guarantee a high data integrity, and offer a high data transparency.

4. Discussion

The proposed model enables a classification of SCEs regarding their transferability to blockchain technology. According to Strong et al. [23], it is essential that an information system consists of a high Contextual DQ and that the stored data provide an added value for the respective use case. Due to the scalability problems of blockchain technology described in Section 3.2, the Contextual DQ becomes even more important, as data without benefit for the respective use case would burden the overall system with unnecessary data traffic. For this reason it is recommended, to first of all filter the SCEs with high contextual DQ when applying this model to a certain use case (1 x x). As Figure 4 indicates, the top layer of the Blockchain Data Quality Cube represents such SCEs. In a second step, the selected SCEs with high relevance can be evaluated regarding their Intrinsic DQ. As described in Chapter 1, the blockchain technology provides a trustworthy, secure, and immutable data storage. Regardless of the Blockchain Suitability DQ, Blockchain projects that mainly consist of SCEs with poor intrinsic DQ (x 3 x) can severely compromise the integrity of the overall system. Therefore, the colour scheme of Figure 4 emphasises all SCEs with a poor Intrinsic DQ. Derived from this, supply chains with low Intrinsic DQ SCEs can be advised to raise the Intrinsic DQ of particularly critical SCEs (for example, by introducing RFID readers or Blockchain Oracles). Finally yet importantly, the selected SCEs can be evaluated regarding their Blockchain Suitability DQ. In particular, SCEs that require low data sizes and infrequent transaction rates are suitable for an on-chain storage. SCEs that do not meet one or more requirements regarding the DQ dimensions described in Section 3.2 are unsuitable for an on-chain storage and must be transferred to a blockchain-verified off-chain storage.

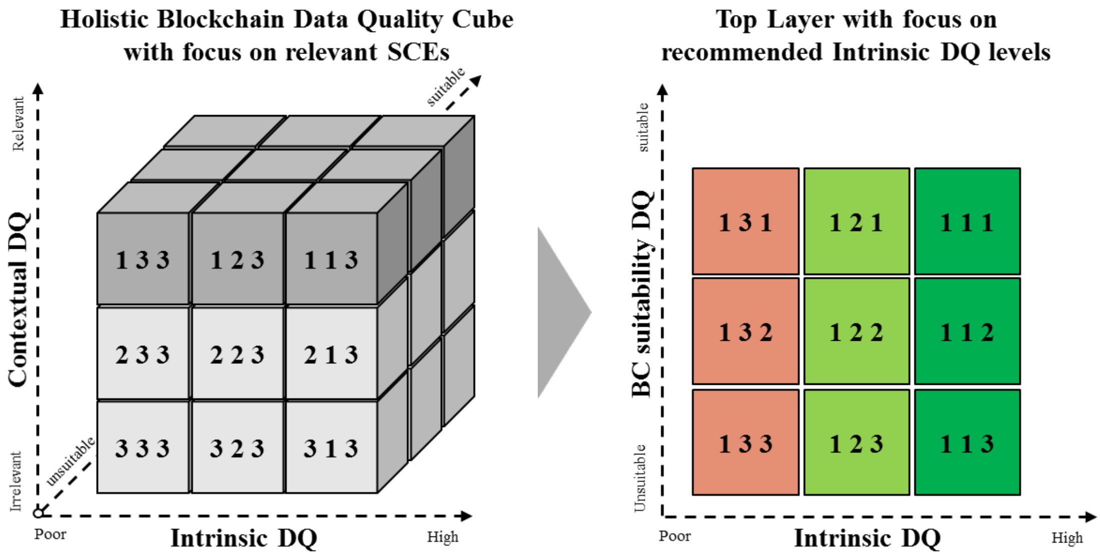


Figure 4: Blockchain Data Quality Cube with evaluation regarding recommended implementation approach

As mentioned in Section 3.1, transferring SCE-Data to an off-chain storage decreases the data transparency of the overall system. Supply chains with many ‘level 2’ Blockchain Suitability DQ Events (x x 2) must be further evaluated regarding the scalability of the overall systems. Although the SCE-Data sets are transferable to the blockchain technology, they could have a negative impact on the usability of the overall system. In such a case a multi-stage evaluation process could be considered, whereby the proposed classification model can be applied again to all SCEs with ‘level 2’ classification (x x 2). These can then be compared among each other and be evaluated based on their Contextual DQ, Intrinsic DQ and Blockchain Suitability DQ.

5. Conclusion

Token-based blockchain architectures enable a holistic mapping of supply chains on the blockchain. Thereby, SCE-Data can either be stored on the blockchain or in blockchain verified off-chain storages. The current limitations of blockchain technology in terms of its scalability have a significant impact when transferring SCEs to the blockchain. Based on the proposed model the authors recommend that the transfer should initially focus only on SCEs with high use case relevance. Here, the intrinsic DQ determines the integrity of the overall system and the Blockchain Suitability DQ determines the data transparency. Blockchain projects that mainly consist of SCEs with poor intrinsic DQ can severely compromise the integrity of the overall system and therefore, can be considered questionable. Regarding the Blockchain Suitability DQ, in particular SCEs that require low data sizes and an infrequent transaction rate, are suitable for an on-chain storage. Otherwise, the data must be stored off-chain. Thereby the block size and the transaction rate of the selected blockchain platform determine the limitation. SCEs with high resource requirements should rather be stored in blockchain verified off-chain storages. SCEs that require high resources but only occur infrequently, for example, can still be stored on-chain. The highest data transparency and integrity can be achieved in supply chains consisting of a large number of SCEs with blockchain suitable events and a high intrinsic DQ. The proposed classification model serves as a theoretical foundation for evaluating supply chain use cases in terms of their sense and feasibility with the adoption of blockchain technology. Further research is currently being conducted to examine the influence of the individual DQ categories on the feasibility of blockchain projects in an industrial scale. Furthermore, a prototype application is currently being developed to evaluate the practical feasibility of such blockchain-based SCE tracking system. In the proposed model, all DQ categories are assessed equally. Similar to the

Failure Mode and Effects Analysis (FMEA), a multiplication of the individual values can be taken into account. Derived from this, a different weighting of the individual DQ categories can possibly take place and a holistic feasibility framework for applying blockchain technology to supply chains can be developed.

References

[1] Dietrich, F., Ge, Y., Turgut, A., Louw, L., Palm, D., 2021. Review and analysis of blockchain projects in supply chain management. International Conference on Industry 4.0 and Smart Manufacturing (ISM) 2, 724–733.

[2] Da Rosa Righi, R., Alberti, A.M., Singh, M., 2020. Blockchain Technology for Industry 4.0: Secure, Decentralized, Distributed and Trusted Industry Environment. Springer, Singapore.

[3] Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H., 2017. An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends. IEEE International Congress on Big Data 6, 557–564.

[4] Nakamoto, S., 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. https://bitcoin.org/bitcoin.pdf. Accessed 29 March 2021.

[5] Alsunaidi, S.J., Alhaidari, F.A., 2019. A Survey of Consensus Algorithms for Blockchain Technology. IEEE International Conference on Computer and Information Sciences (ICCIS), 1–6.

[6] Buterin, V., 2013. A Next-Generation Smart Contract and Decentralized Application Platform. https://ethereum.org/en/whitepaper/. Accessed 29 March 2021.

[7] Dolgui, A., Ivanov, D., Potryasaev, S., Sokolov, B., Ivanova, M., Werner, F., 2020. Blockchain-oriented dynamic modelling of smart contract design and execution in the supply chain. International Journal of Production Research 58 (7), 2184–2199.

[8] Mohanta, B.K., Panda, S.S., Jena, D., 2018. An Overview of Smart Contract and Use Cases in Blockchain Technology. IEEE International Conference on Computing, Communication and Networking Technologies (ICCCNT) 9, 1–4.

[9] Altarawneh, A., Herschberg, T., Medury, S., Kandah, F., Skjellum, A., 2020. Buterin's Scalability Trilemma viewed through a State-change-based Classification for Common Consensus Algorithms. IEEE Annual Computing and Communication Workshop and Conference (CCWC) 10, 727–736.

[10] Konovalenko, I., Ludwig, A., 2019. Event processing in supply chain management – The status quo and research outlook. Computers in Industry 105, 229–249.

[11] Kennedy, A., Southall, M., Morgan, G., Traub, K., 2016. EPC Information Services (EPCIS) Standard. https://www.gs1.org.

[12] Holtkemper, D., 2020. Blockchain-Applikation für das Supply-Chain-Management: Blockchain application for supply chain management. Apprimus, Aachen.

[13] Hepp, T., Sharinghousen, M., Ehret, P., Schoenhals, A., Gipp, B., 2018. On-chain vs. off-chain storage for supply- and blockchain integration. it - Information Technology 60, 283–291.

[14] Abeyratne, S.A., Monfared, R.P., 2016. Blockchain ready manufacturing supply chain using distributed ledger. International Journal of Research in Engineering and Technology 05 (09), 1–10.

[15] Primavera de Filippi, Benedikt Schuppli, Constance Choi, Carla Reyes, et al., Don Tapscott, 2019. Regulatory Framework for Token Sales: An Overview of Relevant Laws and Regulations in Different Jurisdictions Coalition of Automated Legal Applications. Blockchain Research Institute.

[16] Dowling, M. Is non-fungible token pricing driven by cryptocurrencies? Finance Research Letters 2021.

[17] Westerkamp, M., Victor, F., Küpper, A., 2019. Tracing manufacturing processes using blockchain-based token compositions. Digital Communications and Networks.

[18] Westerkamp, M., Friedhelm, V., Kuepper, A., 2018. Blockchain-based Supply Chain Traceability: Token Recipes model Manufacturing Processes. IEEE International Conference on Blockchain.

[19] Watanabe, H., Ishida, T., Ohashi, S., Fujimura, S., Nakadaira, A., Hidaka, K., Kishigami, J., 2019. Enhancing Blockchain Traceability with DAG-Based Tokens. IEEE International Conference on Blockchain.

[20] Dietrich, F., Turgut, A., Louw, L., Palm, D., 2020. Token-Based Blockchain Solutions for Supply Chain Strategies. Production at the leading edge of technology, 689–698.

[21] Antonopoulos, A., Wood, G., 2018. Mastering Ethereum: Building Smart Contracts and Dapps. O'Reilly, Sebastopol.

[22] Eberhardt, J., Tai, S., 2017. On or Off the Blockchain? Insights on Off-Chaining Computation and Data. European Conference on Service-Oriented and Cloud Computing (ESOCC) 6, 3–15.

[23] Strong, D.M., Lee, Y.W., Wang, R.Y., 1997. Data quality in context. Commun. ACM 40 (5), 103–110.

[24] Zheng, Q., Li, Y., Chen, P., Dong, X., 2018. An Innovative IPFS-Based Storage Model for Blockchain. IEEE/WIC/ACM International Conference on Web Intelligence (WI), 704–708.

[25] Kilic, B., Ozturan, C., Sen, A., 2020. A Cluster Based System for Analyzing Ethereum Blockchain Transaction Data. International Conference on Blockchain Computing and Applications (BCCA) 2, 59–65.

[26] Lo, S.K., Xu, X., Staples, M., Yao, L., 2020. Reliability analysis for blockchain oracles. Computers & Electrical Engineering 83.

[27] Mühlberger, R., Bachhofner, S., Ferrer, E.C., Di Ciccio, C., Weber, I., Wöhrer, M., Zdun, U., 2020. Foundational Oracle Patterns: Connecting Blockchain to the Off-chain World. International Conference on Business Process Management, 35–51.

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.

Eldin Mehmedovic (*1993) is currently pursuing a M.Eng. and M.Sc. double degree at Reutlingen University, Germany and Stellenbosch University, South Africa. His main research interests include blockchain applications in supply chain management in combination with Cyber physical Systems and IoT. In 2020, he received his B.Eng. from the Reutlingen University, Germany.

Ali Turgut (*1984) studied industrial engineering (B.Eng.) with a focus on supply chain management at the Esslingen University of Applied Sciences, Germany and the Kettering University, Michigan, USA. He is an employee in the area of innovation management at Steinbeis Innovation gGmbH and is completing a part-time master's degree in Digital Business Engineering (M.Sc.) at the Reutlingen University, Germany.

Louis Louw (*1974) has been an associate professor at Stellenbosch University since 2018. Previously Prof. Dr.-Ing. 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 Value-Added 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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Investigation Of Suitable Methods For An Early Classification On Time Series In Radial-Axial Ring Rolling

Simon Fahle[1], Thomas Glaser[1], Bernd Kuhlenkötter[1]

[1]*Chair of Productionsystems/Ruhr-University Bochum, Bochum, Germany*

Abstract

To increase competitiveness in the hot forming sector, there is a constant urge to improve the rolling process and its products. Industry 4.0 and its impact on data acquisition and data availability enable data driven methods for optimization. In order to optimize the quality prediction of rolled rings in Radial-Axial Ring Rolling (RARR) with regard to ovality as early as possible and hence prevent scrap and unnecessary rework, machine learning methods from the early classification on time series subdomain are used and evaluated within this research. Different approaches from the time series classification domain within supervised learning are used and compared. A so-called minimum prediction length of the ring rolling process time series is analysed using real world production data from thyssenkrupp rothe erde Germany GmbH. Building upon results of earlier research regarding the use of time series classification in RARR by FAHLE ET AL. fully automated as well as domain specific minimum prediction lengths will be investigated. The results of both approaches are compared and evaluated with regards to the current maximum prediction accuracy using the whole sequences, which should provide the highest score as it holds all available information of each sample.

Keywords

Radial-Axial Ring Rolling; TSC; ECTS; Quality

1. Introduction

Preventing scrap and unnecessary rework is a difficult task in many areas of manufacturing. In Radial-Axial Ring Rolling (RARR), production times vary depending on the rolled geometries. Smaller rings with an outer diameter smaller than 1.5 m may need production times less than one minute on the actual ring rolling machine and can be rolled within one heat. Rings with diameters of up to 16 m however need several minutes divided in different heats. For this process, current RARR machines log several channels during the forming and give access to high dimensional information including forming forces, torques, geometric features within the forming process and many more. All those information can be interpreted as a fixed sequence for each rolled ring and thus can be analysed using methods from machine learning. Every sample (representing a single rolled ring) used within this research was measured after the production process using an automated 3D-laser measurement unit, producing accurate targets regarding the ovality of each ring. This combination of logged channels (features) and measured ovality (target) allows the utilization of supervised machine learning methods. Within supervised machine learning, there is the domain of time series classification (TSC), where time series such as the logged channels are used to predict a continuous or discrete label. For the underlying task of predicting ring quality, a discrete label approach is chosen. A general introduction and brief definition of TSC and time series is given.

According to FAWAZ ET AL. and specifications given by FAHLE ET AL. regarding RARR the TSC-task can be formulated as followed: An univariate time series

$$X = [x_1, x_2, \dots, x_T] \quad (1)$$

is a timestamp ordered set of real values with a fixed sampling rate. The length of X is equal to the real values T depending on the individual rolling time of each ring. Each univariate time series represents a logged channel of the RARR process. An M-dimensional multivariate time series

$$X = [X^1, X^2, \dots, X^M] \quad (2)$$

consists of M different univariate times series with $X^i \in \mathbb{R}^T$. A dataset

$$D = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\} \quad (3)$$

is a collection of corresponding pairs (X_i, Y_i) with X_i either being a multivariate or univariate time series and Y_i as its target variable vector. For now, the classes are defined as a binary classification task of either "no ovality" or "ovality" [1,2]. This data set can then be used for a supervised learning task to find a model that learns a function that maps the input sequences to a class label [3]. For more in depth information, a great introduction to the different problem definitions using machine learning in the domain of sequence data and its variants is given by LONING ET AL. [4].

A working TSC model for the application in RARR does provide several advantages such as a cost-efficient post-process labelling of product quality regarding the used threshold for its training data. If the model is sufficiently accurate and precise, this can lead to an alternative way of predicting quality instead of measuring it for every single sample whilst only measuring those that were predicted to be of bad quality and thus save production and energy costs. Other benefits of such a TSC model are several techniques such as a class-activation-map (CAM) which visualizes the decision of the model to show what neuron is firing given each time step within the sequence [2]. However, TSC holds one negative effect, which is the post-process prediction of each label, i.e. the rolling process has to be finished before a prediction of the quality (ovality) can be made by the used model. This does only allow for a prediction and not for a prevention of the regarded quality aspects. Extending the model to prevent bad quality, a prescriptive approach has to be implemented. This can be done by using models and methods of a subdomain called early classification on time series (ECTS) that seeks to classify time series as early as possible while not neglecting the prediction accuracy. In contrast to the approach of labelling a full time series with a label, there is the challenge of classifying a time series as early as possible. Thus, this approach tries to classify time series with less time steps than the TSC approach and therefore is generally a more difficult task to provide sufficient or equally accurate results according to MORI ET AL. [5]. A more in depth and detailed definition of the ECTS approach is given by XING ET AL. and MORI ET AL. [5,6].

In the following section, a short summary of the basic domains i.e. RARR, TSC and ECTS is given by presenting related works in form of current research and industrial advances in each domain.

2. Related Work

The following section covers the related work in the domain of RARR and the two subdomains of the machine learning domain: TSC and ECTS. Current research and state-of-the-art algorithms will be discussed and presented to give a broader overview of the research focus within this experiment.

2.1 Radial-Axial Ring Rolling

Many approaches concerning current research within the domain of RARR are dealing with FEM-Simulation, quality issues and process improvements. For example ALLEGRI ET AL. proposed new milling curves that achieved good quality results with reduced process time and loads based on a FE-Simulation [7].

An innovative control strategy for titanium alloy based on an intelligent FE-simulation is introduced by LI ET AL. to prevent damage to the microstructure of the titanium alloy by controlling its temperature [8]. Enhancing this approach, LIANG ET AL. used an intelligent FE-simulation to plan rolling paths for the ring rolling process of titanium alloy [9].

Process improvements and new and innovative ways to combine RARR with other processes are researched as KUHLENKÖTTER ET AL. investigated the suitability to combine the RARR process with thermal spraying to compact coatings in an innovative manner. Yet a final intact coating could not be established in the initial experiments, promising results, showing that by the subsequent rolling of thermally sprayed coatings the porosity can be significantly reduced and higher compressed residual stresses can be induced, are presented [10]. Another approach where RARR was combined with a different process is shown by MICHL ET AL. by using additive manufacturing in form of a wirearc additive manufacturing. The authors explored whether wirearc additive manufacturing is a suitable solution to produce better near-net shape pre-forms. Referring to their results, wirearc additive manufacturing shows promising results to lower process expenses and improve process efficiency [11].

An overview of the advantages of data driven techniques in metal forming as well as remaining research questions are given by HAVINGA ET AL. such as model uncertainty and measurement accuracy in metal forming [12]. Research regarding machine learning for quality prediction was conducted by different authors. GIORLEO ET AL. used a regression model based on process parameters from FEM data to classify product quality with regards to the fishtail defect [13]. A different form error prediction using machine learning was researched by FAHLE ET AL. regarding ovality. The authors investigated the data utilization and data analysis methods in RARR in 2019 for a variety of German ring rolling companies [14]. Moreover, they used a TSC to classify real world production data to improve the process quality and focused on different preprocessing approaches [1]. This TSC approach will be enhanced throughout this study and the current related work for the domain of TSC and ECTS will be lined out.

2.2 TSC

During recent years, many models and algorithms of different types have been proposed and a comparable data base for validated comparisons was established by introducing the UEA & UCR Time Series Classification Repository [15]. This database was the baseline to validate a variety of models such as HIVE-COTE: a big ensemble with great results on the repository [16]. A big effort was taken by BAGNAL ET AL. in their review of current advances regarding TSC algorithms where a variety of current algorithms was evaluated and compared [17]. Nevertheless, deep learning approaches, which have gained great popularity during the last decade due to their advances in image recognition and other areas, have not been the focus of the review. However, several deep learning approaches to solve TSC tasks have been proposed such as InceptionTime that bases on an Inception architecture [18] or another approach combining the convolution based structure with long short-term memory cells in a so-called MLSTM-FCN network architecture [19]. During the last years several approaches have been taken to form unified architectures for the TSC task, one of those is sktime by LÖNING ET AL. which is compatible with the famous scikit-learn python library and presents a unified interface to deal with time series [4].

2.3 ECTS

SANTOS ET KERN investigated different approaches for the early classification on time series and many agreed on the definition that a classification decision should be made as early as possible while the classification accuracy should not be sacrificed too much. They further investigated different approaches and already included deep learning approaches to solve the ECTS problem [20]. Another recent systematic review for ECTS was published and reviewed different ECTS approaches for many different domains such as Human Activity Recognition, Industrial Process Mining, Quality Monitoring and others [21]. A fairly new approach is proposed by SCHÄFER ET LESER named TEASER. TEASER uses a two-tier approach where a slave-master structure is established to find the best individual decision times for an early classification. The slave classifier classifies the time series and computes class probabilities which are then passed to a master classifier that decides whether the probability is sufficient to emit a final classification result [22]. Another approach that focuses more on the multivariate aspect in ECTS is a confidence-based approach by HE ET AL. Within their research they focus on a multivariable and interpretable approach based on interpretable rules mined from the time series and using subsequences [23]. Moreover, ECTS is used in a variety of different domains. It is used by HATAMI ET CHIRA to classify odor in the bio-chemical domain. They use an ensemble to agree on either accepting or rejecting the class label. The authors use their implementation as the core mechanic for their online e-nose system for odor classification [24]. Moreover, it is applied to other real world datasets within the UCI machine learning repository such as the hydraulic system monitoring data set. One example is the divide-and-conquer-based approach by GUPTA ET AL. [25] or other applications of ECTS such as semiconductor manufacturing [26,27].

3. Experimental setup

The following section describes the essential parts of the later conducted experiments namely the data set all experiments are based on, the preprocessing steps and the used models and approaches.

3.1 Data set

The used data set for this research is provided by thyssenkrupp rothe erde Germany GmbH and is part of its real world production data of different production days. The used samples represent a mixture of different geometric preforms and rolled shapes. New production data is constantly added and preprocessed but the state of the current data set consists of 1078 multivariate samples. Due to a non-disclosure agreement the dataset must not be made available to public. The data set is almost evenly distributed as the threshold for the quality prediction was set to in internal standard to enable a working model to increase production quality by a sufficient amount. The data is acquired at thyssenkrupp together with their representing targets acquired by an independent measurement unit. The multivariate channels/features of the dataset range from radial and axial forces, ring growth rates, motor current and several geometric features of the process. The used data set is split up into a train and test split using an 80/20 split ratio to prevent information leakage. Moreover, a five-fold shuffle split validation is performed for more reliable classification results on the domain specific approach.

All data samples were backpadded to ensure an equal length as process variations lead to different length time series. Backpadding was used because of the requirement that this model is intended to be deployed in a real world radial-axial ring rolling mill system and predicting ovality as early as possible to enable counter measurements. Thus, a backpadded approach is used because the incoming time series length of the process is not known prior due to the different preforms and rolled geometries. Moreover, a second preprocessing is used to the extent that a specific logged channel represents the current rolling phase that the process is in. This channel was used in earlier research by FAHLE ET AL. [1] to investigate suitable preprocessing methods for the application in RARR. Figure 1 depicts an idealized trend of the ring rolling process and its respective rolling phase. The preprocessing that is done in this research in addition to the single backpadding is a cutting between phases, meaning that all samples were separated in four different data sets each containing the samples up to the different rolling phases (e.g. Start,…, End Phase 1; Start, …, End Phase 2; etc.). This is done to investigate a minimum predictive length for the ECTS approach in combination with necessary process specific counter measurements that need to be performed at specific times in the process as they cannot be implemented in every step of the process.

In addition, all samples, neglecting their earlier preprocessing, were examined using standardization, with a scaler fitted on training samples only to ensure no information leakage into the test set, or their raw input values with no standardization. All these preprocessing steps are performed to ensure an early classification as well as ensure system specific requirements due to the process nature.

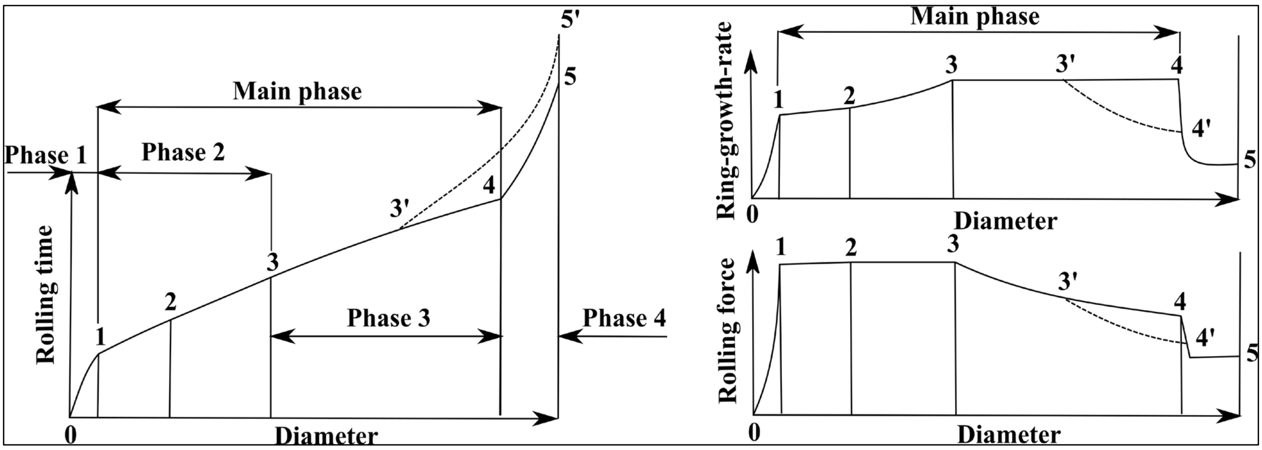


Figure 1: Representation of an idealized rolling phase of a RARR process adapted from [28]

3.2 Baseline ECTS approach

As a baseline the implementation of a non-myopic algorithm is adapted from [29,30]. The algorithm is non-myopic in the sense, that it calculates an optimal time to give a prediction for each time step. This algorithm is chosen as a baseline comparison as it uses a k-nearest-neighbor classifier and chooses the time steps at which a prediction is made automatically, depending on some hyper-parameter choices (used choices in brackets). One parameter was the cluster size (2) for a needed clustering within the model. The next was a lambda (15) value for the evaluation of a prediction during the computation of the cost function. The cost time parameter, alpha (0.0001), is used to shift the trade-off between earliness and accuracy. Lastly, the minimum length (15) for a prediction is set via a separate hyper-parameter. The MPL for the baseline approach needs to be investigated and is only influenced by the hyper-parameter choice.

3.3 Domain-specific rolling phase approach

The domain specific rolling approach has been briefly discussed earlier. The general idea is to predict ovality on-line at designated process states to ensure, that if a prediction is made that requires a process intervention, suitable counter methods can still be applied and the process is not too advanced. This is done even though the accuracy might suffer in order to reach hard earliness requirements. Counter methods such as a dislocation of the ring or an induced change in the axial frame force might not be established with measurable effects if the process is too advanced. This leads to the proposed approach to predict at the end of each rolling phase, depicted in Figure 1. Thus, the MPL for the domain specific approach is set to either one of the four phase endings.

3.4 Summary of chosen approaches

This baseline approach was chosen to investigate the different prediction times, depending on the aforementioned hyper-parameters, if they were chosen automatically. Whereas the domain specific prediction times of the domain specific approach are fixed with regard to each individual process. These prediction times will be investigated and discussed within the evaluation in the next section. It will be analysed whether the domain specific approach and the corresponding process phase endings show similar MPL or are completely different.

4. Experiments

The conducted experiments where divided into two parts. In the first part the automatically chosen prediction times using the baseline approach were investigated using model trainings on the train and test split. The second part uses a domain specific approach on the five-fold shuffle split validation using a Time Series Forest (100 estimators, criterion=entropy, max features=log2), InceptionTime (batch size=32, number filters=64, use_residual=True, use_bottleneck=True, depth=9, kernel size=64) and LSTM-FCN (LSTM=2, Filters=[64,128,64], Kernels=[8,5,3]) model as those models have been proven to be suitable for the time series classification task in the domain of RARR.

4.1 Baseline-Approach

The non-myopic model was used to generate different prediction times. The hyper-parameters were chosen using a grid search and the final parameters are mentioned in section 3.2. The model was trained on both the raw data set as well as on the standardized data set to see if prediction times could vary with respect to value changes due to standardization and the cost time parameter used to force an earlier or later prediction with an accuracy trade-off. It can be seen in Figure 2 that the standardization does change the prediction time for the non-myopic model significantly, as there is a shift from a heavily separated prediction split from either very early and very late prediction towards a more divided prediction, yet the high predictions at a very early stage of the process still occur in the standardized data.

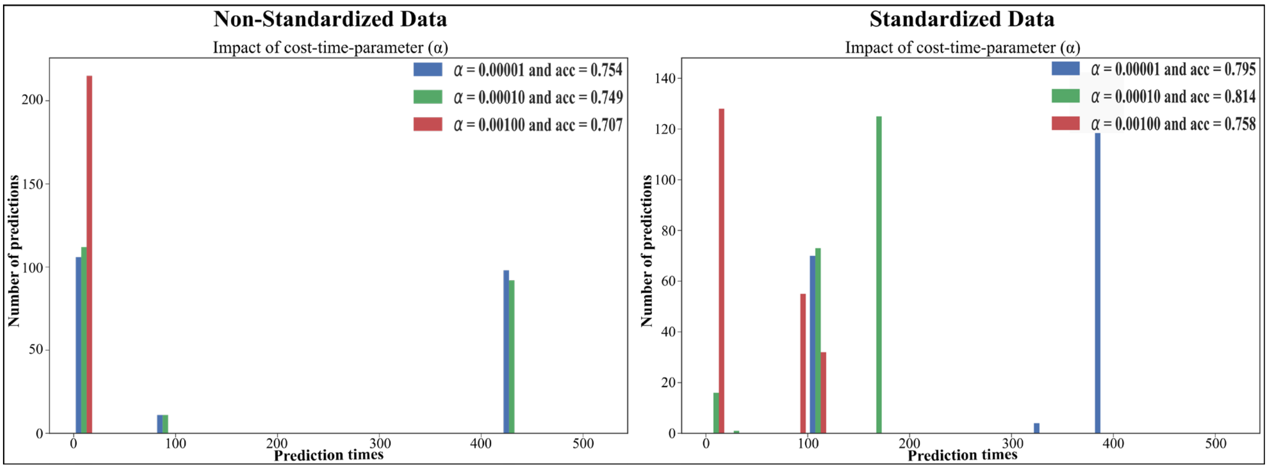


Figure 2: Differences in the earliness of predictions regarding standardization

Moreover, regarding the accuracy within both data sets, it can be seen that with a lower alpha value the accuracy is lower due to the bigger focus on earlier predictions. Yet, it is surprising, that the lowest alpha value does not always produce the highest accuracies, as for the standardized data set predictions with an alpha value of 0.0001 produce a higher accuracy even though the predictions were made earlier than those with alpha 0.00001. Further, Figure 3 depicts two samples (left and right) of the data set that are predicted using the different alpha models and showing the results regarding accuracy. In the example, it can be seen that the prediction made on the bottom (red line) prediction is incorrect for both samples as the time series could not be distinguished correctly by the model at this early stage. The blue (first) and green (second) lines predictions show the later predictions. In this example the latest prediction correctly classifies both example, whereas the second model using an alpha value of 0.0001 does only correctly classify one of the two examples.

Inspecting not only those two examples, it was hard to differentiate whether there is a specific point in the process to which extend all information needed was gathered by the non-myopic approach. Yet, the prediction switches towards the depicted times at about 180 time steps (green) and 390 (blue) seem similar to the process phase approach discussed in the next section.

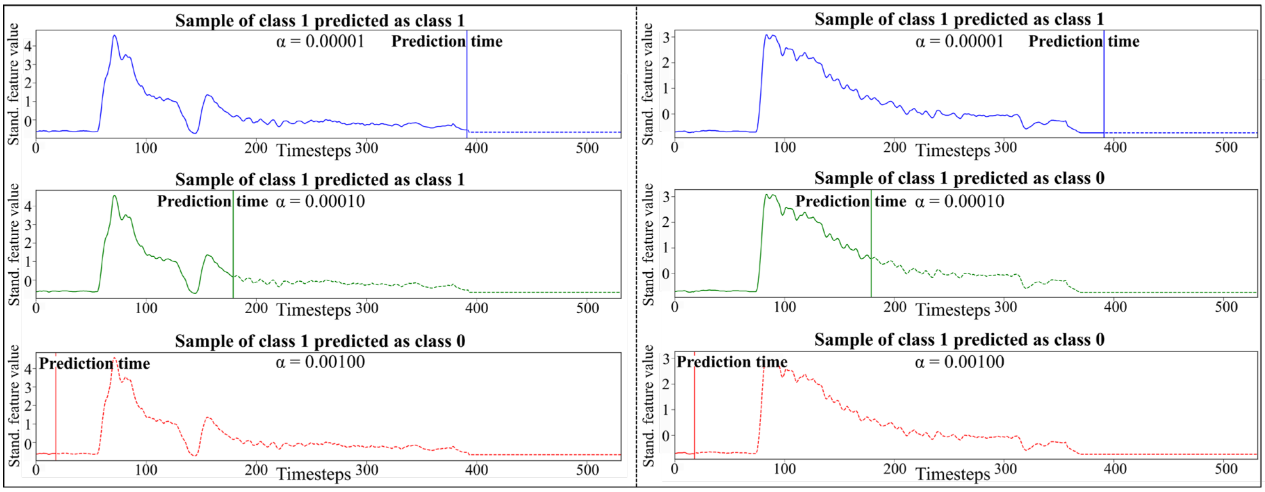


Figure 3: Visualization of different predictions on two samples made by the three different non-myopic models

4.2 Domain specific phase approach

The aforementioned similarity can be seen by analyzing the average phase lengths of each sample. Figure 4 shows the range of the phase lengths on the left side and the cumulated time steps each phase ends on the right. It can be seen that the ending of phase one at about ten time steps is similar to the very early predictions of the non-myopic model. The average phase ending of phase two on the right shows a direct similarity to the model depicted in green in Figure 2. Yet, the average phase ending of phase three at about 320 on the right of Figure 3 is slightly earlier than the prediction made in blue in Figure 2.

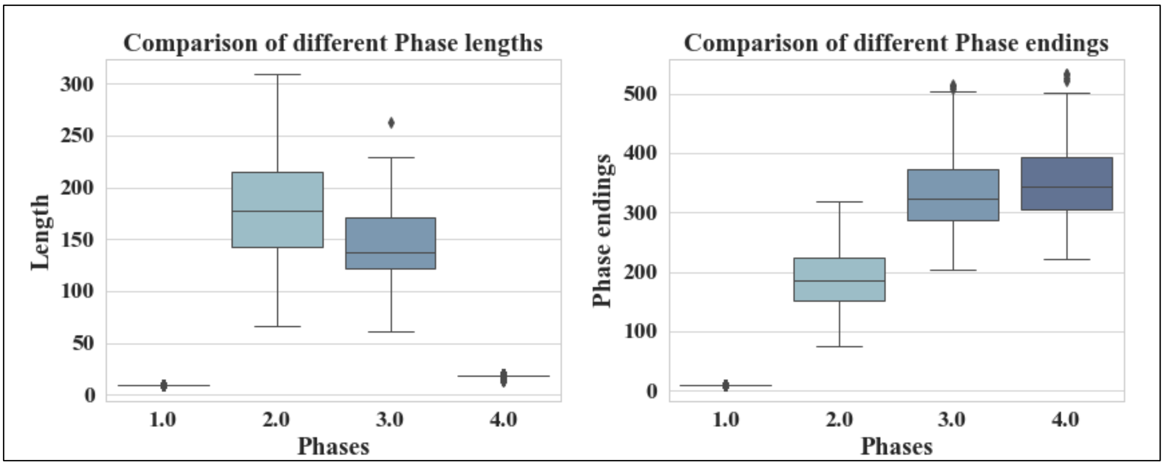


Figure 4: Comparison on the phase lengths and phase endings for the underlying RARR data set

Table 1 shows the test results of the five-fold shuffle split validation training with five runs each for the LSTM-FCN, Inception and Time Series Forest model. It can be seen that the accuracy drop between the predictions after phase one and two as well as phase two and three are noticeable yet between phase three and four are only minor. It is surprising, that the best mean (0.884) and maximum (0.921) accuracy for not standardized data reached for all phases was after phase three and not after phase four. Moreover, it can be seen, that the Time Series Forest performs very well overall and even best when looking at not standardized data in Table 1. For a very early prediction e.g. after phase two the LSTM-FCN Model is performing best.

With regard to the real-world deployment of those models it has to be discussed whether a prediction after phase two or three needs to be forced if the chosen counter measurements need to be applied during the earlier stages of the ring rolling process. The accuracy of the best performing models with regard to the phase is decently high, yet to be deployed into an industrial production chain the accuracies should improve to higher levels.

Table 1: Test-results for the early prediction using the domain specific rolling phase approach

Rolling Phase	No Standardization		Standardization		Classifier
	mean	max	mean	max	
1	0.762	0.805	0.762	0.800	Inception
	0.771	0.800	**0.772**	**0.805**	LSTM
	0.812	**0.847**	0.732	0.753	TSF
2	0.864	0.888	0.856	0.893	Inception
	0.866	**0.902**	**0.867**	**0.907**	LSTM
	0.863	0.888	0.845	0.888	TSF
3	0.868	0.898	0.866	0.898	Inception
	0.871	0.907	0.868	0.898	LSTM
	0.884	**0.921**	**0.878**	**0.907**	TSF
4	0.870	0.907	0.867	0.907	Inception
	0.868	0.907	0.873	**0.912**	LSTM
	0.882	**0.916**	**0.879**	0.907	TSF

5. Conclusion

Within the present research, after reviewing the state of the art for TSC and ECTS, the fundamentals of RARR were explained. Then, the dataset used was described and a baseline and domain-specific approach was presented. Within the experiment section, both approaches were analyzed and compared and achievable accuracies of the ECTS case were presented. The investigation of the different prediction times chosen by the non-myopic approach showed, that the prediction times occur roughly at the same time as the ring rolling process phase changes. This underlines the used approach to link the prediction of different models to the rolling phase changes as it is done in section 4.2. The overall better performance of the domain specific approach is due to the different and more complex models used and the baseline approach was used to gather useful information about the minimum prediction length of the data. For the implementation at the radial-axial ring rolling mill at the chair of production system real world experiments will be carried out using the early prediction models to initiate counter measurements to avoid ovality, resulting in better quality.

6. Future Work

The aforementioned results will be used to deploy an on-line system for quality prediction at a ring rolling mill. Depending on the different levels of earliness of the predictions different counter measurements will be triggered via a direct connection to the CNC: Further validations regarding those counter measurements will be conducted. Moreover, the earliness and accuracy will be constantly improved and the data set will be enlarged to achieve higher prediction accuracies throughout all rolling phases.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 404517758

The authors would like to additionally thank Dr.-Ing. Andreas Kneißler and thyssenkrupp rothe erde Germany GmbH for their support and access to their real world data.

References

[1] Fahle, S., Kneißler, A., Glaser, T., Kuhlenkötter, B., 2021. Research on Preprocessing Methods for Time Series Classification Using Machine Learning Models in the Domain of Radial-Axial Ring Rolling, in: Behrens, B.-A., Brosius, A., Hintze, W., Ihlenfeldt, S., Wulfsberg, J.P. (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 487–496.

[2] Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., Muller, P.-A., 2019. Deep learning for time series classification: a review. Data Min Knowl Disc 59 (2), 195.

[3] Xing, Z., Pei, J., Keogh, E., 2010. Xing et al. A Brief Survey on Sequence Classification. SIGKDD Explorations (Volume 12, Issue 1), 40–48.

[4] Löning, M., Bagnall, A., Ganesh, S., Kazakov, V., Lines, J., Király, F.J., 2019. sktime: A Unified Interface for Machine Learning with Time Series (Preprint), 1–9.

[5] Mori, U., Mendiburu, A., Keogh, E., Lozano, J.A., 2017. Reliable early classification of time series based on discriminating the classes over time. Data Min Knowl Disc 31 (1), 233–263.

[6] Xing, Z., Pei, J., and Philip, S. Y., 2012. Early classification on time series. Knowledge and information systems (31 (1)), 105–127.

[7] Allegri, G., Giorleo, L., 2020. Ring Rolling speed rolls optimization to improve ring quality and reduce production time. JMES 14 (1), 6272–6284.

[8] Li, X., Guo, L., Wang, Y., Liang, L., 2019. Rolling paths design assisted by target-temperature driven intelligent FE simulation of radial-axial ring rolling. Manufacturing Rev. 6, 20.

[9] Liang, L., Guo, L., Wang, Y., Li, X., 2019. Towards an intelligent FE simulation for real-time temperature-controlled radial-axial ring rolling process. Journal of Manufacturing Processes 48, 1–11.

[10] Kuhlenkötter, B., Glaser, T., Fahle, S., Husmann, S., Abdulgader, M., Tillmann, W., 2020. Investigation of compaction by ring rolling on thermal sprayed coatings. Procedia Manufacturing 50, 192–198.

[11] Michl, D., Sydow, B., Bambach, M., 2020. Ring Rolling of Pre-forms Made by Wire-arc Additive Manufacturing. Procedia Manufacturing 47, 342–348.

[12] Havinga, J., Mandal, P.K., van den Boogaard, T., 2019. Exploiting data in smart factories: real-time state estimation and model improvement in metal forming mass production. Int J Mater Form 207 (3), 663–673.

[13] Giorleo, L., Ceretti, E., Giardini, C. Investigation of the fishtail defect in ring rolling by a FEM approach, in: , North American Manufacturing Research Institution, North American Manufacturing Research Conference et al. 2012, vol. 40, pp 104–111.

[14] Fahle, S., Kuhlenkötter, B., 2020. A Framework for Data Integration and Analysis in Radial-Axial Ring Rolling, in: Nyhuis, P., Herberger, D., Hübner, M. (Eds.), Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020). Hannover : Institutionelles Repositorium der Leibniz Universität Hannover, 127-136.

[15] Dau, H.A., Keogh, E., Kamgar, K., Yeh, C.-C.M., Zhu, Y., Gharghabi, S., Ratanamahatana, C.A., Yanping, Hu, B., Begum, N., Bagnall, A., Mueen, A., Batista, G., Hexagon-ML, 2018. The UCR Time Series Classification Archive.

[16] Lines, J., Taylor, S., Bagnall, A., 2018. Time Series Classification with HIVE-COTE. ACM Trans. Knowl. Discov. Data 12 (5), 1–35.

[17] Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E., 2017. The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min Knowl Disc 31 (3), 606–660.

[18] Fawaz, I.H., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D.F., Weber, J., Webb, G.I., Idoumghar, L., Muller, P.-A., Petitjean, F., 2020. InceptionTime: Finding AlexNet for time series classification. Data Min Knowl Disc 34 (6), 1936–1962.

[19] Karim, F., Majumdar, S., Darabi, H., Harford, S., 2019. Multivariate LSTM-FCNs for time series classification. Neural networks : the official journal of the International Neural Network Society 116, 237–245.

[20] Santos, T., Kern, R., 2016. A Literature Survey of Early Time Series Classification and Deep Learning.

[21] Gupta, A., Gupta, H.P., Biswas, B., Dutta, T., 2020. Approaches and Applications of Early Classification of Time Series: A Review.

[22] Schäfer, P., Leser, U., 2020. TEASER: early and accurate time series classification. Data Min Knowl Disc 34 (5), 1336–1362.

[23] He, G., Zhao, W., Xia, X., 2020. Confidence-based early classification of multivariate time series with multiple interpretable rules. Pattern Anal Applic 23 (2), 567–580.

[24] Hatami, N., Chira, C., 042013. Classifiers with a reject option for early time-series classification, in: 2013 IEEE Symposium on Computational Intelligence and Ensemble Learning (CIEL). 2013 IEEE Symposium on Computational Intelligence and Ensemble Learning (CIEL), Singapore, Singapore. 16.04.2013 - 19.04.2013. IEEE, pp. 9–16.

[25] Gupta, A., Gupta, H.P., Biswas, B., Dutta, T., 2020. A Divide-and-Conquer–based Early Classification Approach for Multivariate Time Series with Different Sampling Rate Components in IoT. ACM Trans. Internet Things 1 (2), 1–21.

[26] He, G., Duan, Y., Peng, R., Jing, X., Qian, T., Wang, L., 2015. Early classification on multivariate time series. Neurocomputing 149, 777–787.

[27] Lin, Y.-F., Chen, H.-H., Tseng, V.S., Pei, J., 2015. Reliable Early Classification on Multivariate Time Series with Numerical and Categorical Attributes, in: Cao, T., Lim, E.-P., Zhou, Z.-H., Ho, T.-B., Cheung, D., Motoda, H. (Eds.), Advances in Knowledge Discovery and Data Mining, vol. 9077. Springer International Publishing, Cham, pp. 199–211.

[28] Thyssen, 1990. Technologiehandbuch: Einführung in die Ringwalztechnologie. Thyssen Wagner Maschinenbau GmbH Dortmund.

[29] Dachraoui, A., Bondu, A., Cornuéjols, A., 2015. Early Classification of Time Series as a Non Myopic Sequential Decision Making Problem, in: Appice, A., Rodrigues, P.P., Santos Costa, V., Soares, C., Gama, J., Jorge, A. (Eds.), Machine Learning and Knowledge Discovery in Databases, vol. 9284. Springer International Publishing, Cham, pp. 433–447.

[30] Tavenard, R., Faouzi, J., Vandewiele, G., Divo, F., Androz, G., Holtz, C., Payne, Marie, Yurchak, Roman, Rußwurm, M., Kolar, K., Woods, E., 2020. Tslearn, A Machine Learning Toolkit for Time Series Data. Journal of Machine Learning Research 21, 118 (21), 1–6.

Biography

Simon Fahle (*1994) is a member of the Lehrstuhl für Produktionssysteme (LPS) at the Ruhr-University Bochum since 2019. He earned a bachelor’s and master’s degree in mechanical engineering at the Ruhr-University Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Thomas Glaser (*1990) is a member of the Lehrstuhl für Produktionssysteme (LPS) of the Ruhr-University Bochum since 2019. He holds a bachelor's as well as a master's degree in mechanical engineering from the Ruhr-University Bochum, specialised in materials engineering. Thomas Glaser is currently focussing on research topics in the field of radial-axial ring rolling.

Until 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 has held the professorship for “Production Systems” at the Ruhr-University Bochum.

2nd Conference on Production Systems and Logistics

Analysis Of Process Models For The Business Model Development Considering Special SME Requirements For Offering PSS

Dominik Lins[1], Dominik Arnold[1], Christian Köhler[2], Tobias Mahl[2], Christopher Prinz[1], Bernd Kuhlenkötter[1]

[1] *Ruhr-University Bochum, Chair of Production Systems, Bochum, Germany*
[2] *Saarland University of Applied Sciences, Saarbrücken, Germany*

Abstract

The continuous testing and redesign of current business models and the associated adaptation of increasingly customized value propositions are nowadays becoming more and more important for companies. Thereby the additional or integrated offering of services as new forms of hybrid value creation are gaining importance. Small and medium-sized enterprises (SMEs) in particular are still at the beginning of the challenge of successfully designing the transformation to offer product service systems (PSS), as they often lack time, technical and methodological resources. This paper first describes the derived requirements for a process model of PSS-oriented business model development for SMEs. The requirements result on the one hand from the research of current scientific publications on this topic and on the other hand from the analysis and evaluation of use cases from practice at SMEs. This is followed by an overview of current process models that deal with business model development and innovation. The identified process models are analyzed in terms of the considered phases, the used methods, and the industries of the potential users. Furthermore, they are examined for SME suitability on the basis of the derived requirements. By means of the comparison of the current process models and the requirements from the SMEs the necessity of an adaption and further detailing of the individual phases of the process models is pointed out. SMEs have to be able to apply these models independently. Finally, the article presents an initial approach specifying the development phase of PSS-oriented business models by means of detailed planning with a focus on the mechanical and plant engineering sector to support the users.

Keywords

hybrid value creation; business model development; process planning; product service systems; SME

1. Introduction

After the degree of digitization of SMEs has continuously increased in recent years, companies are now at an important decision point: as the current IFM study "Future Panel Medium-Sized Businesses 2020" [1] shows, SMEs have acquired a clearer understanding of digitization, In this insight digitization is not a self-contained target, but gives enterprises much more potential. During the first phases of digitization discovered potentials must be used and the skills acquired must be channeled in such a way that new innovative value propositions can be created. A solution approach that has already been discussed for three decades and addresses precisely this integration is the service transformation, i.e. the transformation of manufacturing companies towards providers of product service systems (PSS) [2]. A PSS is an integrated offer of one or more goods and services [3].While many companies and especially SMEs have long shied away from initiating their service transformation, it has recently become apparent that digitization is acting as a driver in this process. By the measurability of usage data, new digital technologies enable, for example, more transparent and controllable service processes as well as innovative billing models [4]. It can be observed that SMEs are also increasingly making this change in strategy of a service transformation. For the development of PSS, a holistic systemic view must always be taken in order to define all relevant elements of a production and service system and their relationships to each other. The transformation process must thus be considered in direct connection with the possible ecological, technological, socio-cultural and institutional potential for change in the production and consumption system. Such a systemic view in the innovation process from classic business models to hybrid business models also creates challenges for developers, especially in SMEs. The paper explores the research question if existing process models for business model transformation meet the requirements for the business model transformation to offer PSS in producing SMEs. Therefore the needed requirements from practice which also can be observed in literature are derived in chapter 2. The literature review for process models and the synthesis of general phases for the transformation process is described in chapter 3. Also the analysis and the evaluation of the selected process models is contained. The last chapter summarizes the results in consideration of the research question and gives an outlook for further research.

2. SME requirements for process models for business model transformation to offer PSS

The Plan-Do-Check-Act (PDCA) cycle is an established model for changing processes, especially in SMEs [5]. With its phase-specific structure, it serves as a best-practice example and a possible basis of comparison for the requirements to be considered for a process model for the transformation of the business model. Therefore, it can be seen as a starting point for any necessary adjustments.

The PDCA cycle consists of four basic phases, which in turn can be divided into a total of six more detailed phases: At the beginning, the task or a problem must first be *identified*. This task is then *analyzed* so that solutions can be *developed*. The developed solutions can then be *implemented* and *evaluated*. Finally, the successful solution to the task must be *standardized* to ensure a sustainable solution to the task. [5] How extensive individual phases of the PDCA cycle are or whether they are complete depends on the requirements of the target group and the target context. Requirements for process models therefore relate not only to the phases themselves but also to the content that should be covered by the existing phases of the models. Thus, in the area of service development or servitization of a product, for example, it is necessary to include the customer as well as the structure of the company-specific processes or to involve employees in a participatory manner in order to avoid resistance to organizational and cultural change [6]. Further requirements for corresponding process models, thus often result from various implementation barriers that have to be overcome during the transformation process.

A transformation of the business model is particularly challenging for SMEs due to limited resources and competencies [7]. Within the ongoing research project, further requirements were derived from expert interviews and from workshops on topics such as lessons learned, customer journey, or stakeholder analyzes based on three different SME use cases. As a result, numerous important aspects were identified. These include, for example, a sufficiently long phase for explaining the topic and recognizing its benefits. In this way, interest in a transformation can be generated among all participants. In addition to interest in the topic, a process model should also generate an understanding of the company's own processes, potential, industries and customers right from the start. Above all, from the point of view of the application companies, a process model should also provide content to check whether certain minimum requirements prevail in the company or include methods to fulfill them. After the requirements that are specifically directed at the first phases of a process model, however, requirements also arose for the development phases. Here, it is necessary for SMEs that a process model contains best practices or concrete examples for inspiration, so that their own creativity is stimulated and they can also think outside the traditional corporate processes. However, this also means that risks can be identified for both the existing business model and the new, innovative business model ideas. Only on the basis of known risks a well-founded decision can be made as to whether the company's own business model should be transformed.

If the planned transformation is also to be implemented, a process model should also cover more specific areas in terms of content and, in relation to the offering of data-oriented PSS, also provide information on how the existing data infrastructure must be adapted or what data infrastructure must be created. Another area that a corresponding process model should cover would be practicable assistance for contracting and pricing the newly developed offerings. The concepts developed within a process model should then also lead to a detailed process concept for implementation in order to get from the initial state to the target state step by step. The required qualifications of key personnel should also be identified. At the end of the process model, methods for anchoring the target state should also be shown.

Table 1: Special content and target requirements of SMEs for a process model for the transformation

Content and target requirements primarily based on SME-use-cases *A process model should ...*
... include a sufficiently long phase for explaining the issue and recognizing its benefits, as well as generating interest in a transformation.
… generate an understanding of the company's own processes, potentials, industries, and customers right from the start.
... check whether certain minimum requirements exist or include methods to fulfil them.
... include best practices or concrete examples for inspiration.
... identify risks of the existing business model and also of the developed concepts.
… include information on how the existing data infrastructure must be adapted or what data infrastructure must be created.
... include workable guidance on contracting and pricing for new offerings.
... include a detailed process concept to get from the initial state to the target state.
... identify the required qualifications of key personnel.
... include methods for anchoring the target state.

In addition to the content requirements for process models summarized, the interviews and workshops already mentioned provided insights into cross-cutting requirements that apply independently of individual phases or content, e.g. how the content of the process model should be communicated or applied. This includes, for example, the approach of first initiating a process top-down via the management level of a company and then shaping the successful implementation bottom-up via the operational level in a creative and participatory manner.

From the perspective of the user of the process model, one of the mode requirements is that, wherever possible, content is conducted or conveyed within personal discussions between experts and SME stakeholders. In this way, employees of transforming companies feel more valued and can receive direct assistance even when problems arise. This also means, for example, that content of the process model should be easy to understand, especially in terms of language, so that employees with different skills and qualifications have the same opportunities to understand it. Content should also be transparent at all times with regard to its significance and its specific effects on the overall process. In this way, SMEs do not get the feeling that they could be wasting their valuable time senselessly. This also leads to the fact that content in the process model could be accomplished temporally as well as financially with low expenditure. In addition, a process model should be implemented with change management methods and a lot of communication between all SME participants at the various company levels. In addition to the basic phase-specific requirement mentioned at the beginning, the mode requirements derived from use cases are summarized in Table 2.

Table 2: Special mode requirements of SMEs for a process model for the transformation

Mode requirements primarily based on SME-use-cases *Contents in the process model should ...*
… first be initiated top down at the management level and then developed and implemented sustainably bottom up at the operational level.
… be carried out or conveyed within personal discussions between experts and SME participants.
… be easy to understand, especially in terms of technical language.
… be transparent in terms of its importance as well as its specific impact in the overall process.
… be implemented at low costs in terms of both, time and money.
… be implemented with change management methods and a lot of communication between all SME stakeholders.

It has been shown in practice that SMEs are interested in a specific process model for PSS business model development, provided it meets their corresponding requirements. In the following, existing process models are therefore analysed and checked for their suitability.

3. Analysis of selected process models for the business model development

This chapter presents the procedure and the actual analysis of selected process models. Starting point was a literature research for process models for the business model development which, among others, also consider the offering of PSS.

3.1 Literature review

First of all, various criteria and boundary conditions for the literature research were defined. The research should be limited to process models that were published from 2010 onwards. This restriction addresses the requirement that explicit methods and technologies should be specified that are of interest for current production systems in consideration of the digitization. In addition, the literature, as mentioned above, should provide a process model with different phases and not only list points and facts that need to be observed. The researched papers were reviewed if they contain or name other process models in their work as foundational or state-of-the-art. The papers thus found are also reviewed against the criteria and, if they met them, selected accordingly for analysis. All in all 25 process models were collected and present the input for the actual analysis.

Before starting the analysis, a synthesis, based on selected process models is carried out. It aims to define general phases for the business model development, because many process models name and categorize their phases differently. The general phases are necessary to compare and jointly analyze the models. For this purpose, the steps of the four well-known process models *Business Model Innovation* [8], *Business Model Management* [9], *Developing Business Models: 55 innovative concepts with the St. Gallen business model navigator* [10] and *Business Model Generation* [11] are compared and the general phases for the analysis derived. In his approach, Schallmo considers the six steps of *idea generation, vision development, prototype development, business model development, business model implementation* and *business model extension* [8]. Wirtz and Daiser draw on a total of seven steps, ranging from *analysis*, *ideation* and *feasibility analysis*, through *prototyping, decision-making* and *implementation* to *sustainability* [9]. The third model by Gassmann et al. specifies the four steps of *initiation*, *idea generation*, *integration* and *implementation* for the development of business models [10]. By developing business models according to Osterwalder and Pigneur, the five steps *mobilize*, *understand*, *design*, *implement* and *execute* are proposed and explained [11]. Based on the steps of these models, the following six general phases are derived for the analysis of the process models: *preparation*, *idea generation*, *design*, *evaluation and selection*, *implementation* and *sustainment*. The first phase *preparation* addresses activities which help to call attention and prepare the company for the start of the transformation process. The phases *idea generation, design* and *evaluation and selection* describe the creative process for new ideas, their design and evaluation. The implementation first in pilot areas and for pilot products and later for the whole company or business organisation happens in the fifth phase while the operation as well as the continuous improvement process of the business model is carried out in the last phase.

3.2 Conduction and evaluation

For the detailed analysis the steps of each researched process model for business model development are assigned to the general phases. During the assignment, each step is transferred to the general phases by noting the heading and a short description or keywords of the step for the corresponding phase (see Table 3). At the same time, it is also analyzed for each of the phases considered whether the process model mentions methods or tools that support the implementation of this step.

Table 3: extract of the analysis sheet

model / phase	preparation	idea generation	design	evaluation & selection	implemen-tation	sustainment
Lins et al. (2021) [6]	attention (1), requirement (2), current status (3)	creative phase (4)	prototyping (5), development (6)	contained in development	implemen-tation (7)	continuity (8)
Osterwalder and Pigneur (2011) [11]	mobilize (1)	understand (2)	design (3)	contained in design	implement (4)	manage (5)
…						

Table 3 shows an extract from the analysis sheet documenting the process models. It becomes clear that the models can all be assigned to the general phases. In some cases, several steps are assigned to one phase (cf. Lins et al. (2021) phase *preparation*) or 2 phases are addressed with one step, so that this step is valid for 2 phases (cf. Osterwalder and Pigneur (2011) phase *evaluation & selection*). In addition to the assignment, the two exemplary models suggest methods that support the implementation of the respective step. The Ability model, for example, mentions a best practice database for creating attention in step (1). Both models use the Business Model Canvas method, among others, to record the current status of the business model and prepare for a successful business model development project (cf. Lins et al. (2021) step (3) and Osterwalder and Pigneur (2011) step (1)). If methods are mentioned by the authors, they are listed in the analysis sheet and assigned to the phases as well as to the steps of the process models. All in all, by analysing the 25 process models all general phases are filled with steps even if some process models only cover a few phases of a complete transformation as evaluated hereafter. The results of the detailed analysis of the individual process models are transferred to a table that provides information on which of the general phases are taken into account in the models (see Table 4). An "X" means that the model includes at least one step that can be classified in the corresponding phase. The six phases are considered with different frequency (see Figure 1).

All models contain steps that can be assigned to the *design* phase, and 22 of the 25 process models consider the phase *idea generation*. Not quite as often the phases *preparation* (14 of 25), *evaluation and selection* (16 of 25), as well as the phase *implementation* (19 of 25) are addressed. The fewest models include process steps that can be assigned to the phase *sustainment*. The evaluation and selection of the business model ideas is already included as a step in the design of several models, but is then marked and highlighted in the analysis for both phases (cf. *Osterwalder and Pigneur* (2011) in Table 3).

Table 4: phases considered in the analyzed process models

model / phase	preparation	idea generation	design	evaluation & selection	implemen-tation	sustainment
Amit und Zott (2015) [12]		X	X			
Amshoff (2016) [13]	X	X	X	X		
Boßlau (2014) [3]			X	X		
Bucherer (2010) [14]	X	X	X	X	X	X
Echterhoff (2018) [15]			X	X	X	
Echterhoff et al. (2017) (*project GEMINI*) [16]	X	X	X	X		
Enkel und Mezger (2013) [17]		X	X		X	
Eurich et al. 2014 [18]		X	X	X	X	
Frankenberger et al. (2013) [19]	X	X	X		X	
Gassmann et al. 2013 [10]	X	X	X		X	
Johnson (2010) [20]	X		X		X	
Köster (2014) [21]		X	X	X	X	
Lehner (2016) [22]	X	X	X	X		
Lins et al. (2021) (*project ABILITY*) [6]	X	X	X	X	X	X
Osterwalder und Pigneur (2011) [11]	X	X	X	X	X	X
Peitz (2015) [23]	X	X	X	X	X	
Pynnönen et al. (2012) [24]	X	X	X	X	X	
Rose (2015) [25]		X	X	X	X	
Schallmo (2013) [8]		X	X		X	X
Sosna et al. (2010) [26]		X	X		X	X
Teece (2010) [27]	X	X	X			
Van der Pijl et al. (2016) [28]	X	X	X	X	X	
Weiner et al. (2010) [29]		X	X		X	
Wirtz (2010) [9]		X	X	X	X	X
Wirtz und Daiser (2018) [30]	X	X	X	X	X	X

The evaluation of the general phases mentioned shows that the process models take into account about four general phases on average. At the same time, according to the evaluation, the individual process models have an average of 4.84 process steps. The most steps are proposed in the models Lins et al. (2021) with eight and Wirtz and Daiser (2018) with seven. The lowest number of steps in the process models is 3 (cf. [12,3,17,20,29]). Regardless of the number of steps, 16 of the 25 models explicitly mention methods and tools for a successful execution of the steps (e.g. [15,20,8]) and three others state at least supporting guiding questions or starting points for the transformation. A useful tool to support the idea generation as well as the elaboration and the design is the application of business model patterns. Five of the analyzed models suggest using patterns and contain possible samples of different business models of different branches (cf. e.g. [16,10]). These can be examples on actual business models as well as samples with abstracted models. Another way of assisting the successful passing through the phases of the process models is to provide a software tool that helps the users to transform their business models. Boßlau (2014) and Echterhoff et al. (2017), for example, provide tools in form of software programs that support the documentation und selection of business model design elements through stored content as well as their links and dependencies.

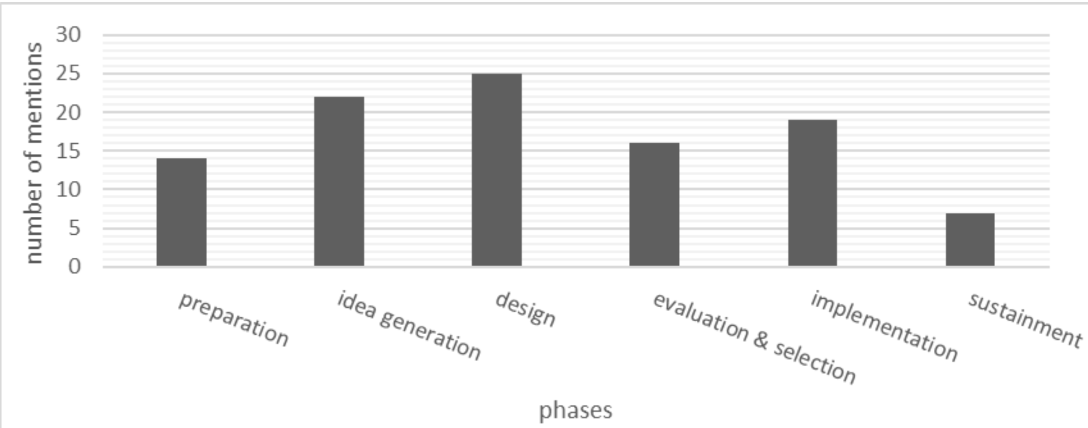


Figure 1: phases considered by the process models

3.3 Summary of the analysis results considering the special requirements

The analyzed process models cover all derived general phases with their respective process steps. Most of the steps can be assigned to the phases *idea generation* and *design*. Starting from these two phases in the transformation process, the number of models that cover the phases with their steps decreases both forwards and backwards. In accordance with the requirements for process models for producing SMEs derived in chapter 2, it can be stated that many models cover a large proportion of the phases for holistic transformation, which is particularly important for SMEs. Pointing out practical methods and tools that can be used in numerous models also supports the use and applicability of the models in manufacturing companies. This is also strengthened by the models that accompany the implementation of the steps of the process model in terms of software programs. Some of the process models have more detailed steps (more than 6 process steps), which partially range from raising awareness and documenting the current business model to operating the new business model (e.g., [6] and [30]). This supports the transformation of SMEs in particular, because the companies can understand the steps well and apply them better. Besides the detailed execution, the business model patterns and best practices are a good help and orientation for users of the process model. A more in-depth detailing of the patterns or of possible characteristics of business models to explicit process descriptions for parts or elements of the new business model in the form of specific process modules is not suggested in any of the analyzed models.

4. Conclusion

The procedure of a business model innovation is fundamentally similar to the procedure of certain improvement processes, whereby the occurrence of the preparation phase and thus ensuring the readiness for transformation is the most important adaptation and is in line with the special requirements of SMEs for offering PSS. The analysis has shown that there is already a large number of process models for business model development and innovation. The models cover the entire transformation process and provide numerous methods and tools, most of which can be used in an industry-neutral manner. However, comparison of the models with the specific requirements of producing SMEs shows that most of the models do not ensure a close support for the manufacturing companies and that there is no specific preconception of individual processes. This poses challenges for SMEs, since they often do not have the technical, human and time resources to work their way through the process models independently, parallel to their daily business. It is assumed that the generation of ideas is sufficiently supported by the known models and that these ideas can be used as input for a detailed planning. The detailed planning is initially to be implemented industry-specifically for the mechanical and plant engineering sector and, in particular, to support SMEs in designing and planning the implementation of new business models for PSS. To this end, standardized processes are to be derived and the necessary elements and resources for the design are to be defined. This detailed planning is intended to enable SMEs to move independently from their actual business model to their target or their desired business model. For a successful transformation of companies in the future, also other topics must be considered in more detail, which are either only touched by the analyzed process models or not considered at all. In addition to detailed planning, these include a risk management and a simulation of the effects of potential risks of the business models to be developed or transformed. Moreover, a user-friendly way of pricing PSS transparently and comprehensibly would meet with great approval among SMEs. Overall, it can be stated that many detailed process models for business model development and innovation exist, but they do not yet fully meet the requirements with regard to support the transformation of SMEs to a business model for PSS. Thus this suggests other approaches for considering specific phases of the transformation process more in detail.

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" programme.

References

[1] IfM Bonn, 2020. Herausforderungen des deutschen Mittelstands in der Corona-Pandemie: Sonderauswertung des Zukunftspanel Mittelstand.

[2] Vandermerve, S., Rada, J., 1988. Servitization of business: Adding value by adding services. European Management Journal 6 (4), 314–324.

[3] Boßlau, M., 2014. Business Model Engineering: Gestaltung und Analyse dynamischer Geschäftsmodelle für industrielle Produkt-Service Systeme. Dissertation, Bochum.

[4] 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.

[5] Johnson, C.N., 2002. The benefits fo PDCA: Use this cycle for continual process improvement. Quality Progress 35 (5), 120.

[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] Schallmo, D., 2013. Geschäftsmodell-Innovation: Grundlagen, bestehende Ansätze, methodisches Vorgehen und B2B-Geschäftsmodelle. Springer Gabler, Wiesbaden, 335 Seiten.

[9] Wirtz, B.W., 2010. Business Model Management: Design - Instrumente - Erfolgsfaktoren von Geschäftsmodellen, 1. Aufl. ed. Gabler, Wiesbaden, 361 pp.

[10] Gassmann, O., Frankenberger, K., Csik, M., 2013. Geschäftsmodelle entwickeln: 55 innovative Konzepte mit dem St. Galler Business Model Navigator. Hanser, München, 320 pp.

[11] Osterwalder, A., Pigneur, Y., 2011. Business Model Generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus Verlag, Frankfurt, New York, 282 pp.

[12] Amit, R., Zott, C., 2015. Creating Value Through Business Model Innovation. MIT Sloan Management Review, 36–44.

[13] Amshoff, B., 2016. Systematik zur musterbasierten Entwicklung technologie-induzierter Geschäftsmodelle. Verlagshaus Monsenstein und Vannerdat OHG, Münster.

[14] Bucherer, E., 2010. Business Model Innovation - Guidelines for a Structured Approach. Shaker Verlag GmbH, Aachen.

[15] Echterhoff, B., 2018. Methodik zur Einführung innovativer Geschäftsmodelle in etablierten Unternehmen. readbox publishing GmbH, Münster, 160 pp.

[16] Echterhoff, B., Koldewey, C., Gausemeier, J., 2017. Pattern based business model development - identification, structuring and application of business model patterns. The ISPIM Innovation Forum.

[17] Enkel, E., Mezger, F., 2013. Imitation processes and their application for business model innovation: An explorative study. International Journal of Innovation Management 17 (01).

[18] Eurich, M., Weiblen, T., Breitenmoser, P., 2014. A Six-Step Approach to Business Model Innovation. International Journal of Entrepreneurship and Innovation Management 18 (4), 330–348.

[19] Frankenberger, K., Weiblen, T., Csik, M., Gassmann, O., 2013. The 4I-framework of business model innovation: a structured view on process phases and challenges. International Journal of Product Development 18, 249–273.

[20] Johnson, M.W., 2010. Seizing the White Space: Business Model Innovation for Growth and Renewal. Harvard Business Press, Boston (Massachusetts), 1 online resource (1 volume).

[21] Köster, O., 2014. Systematik zur Entwicklung von Geschäftsmodellen in der Produktentstehung. Verlagshaus Monsenstein und Vannerdat OHG, Münster.

[22] Lehner, A.-C., 2016. Systematik zur lösungsmusterbasierten Entwicklung von Frugal Innovations. Verlagshaus Monsenstein und Vannerdat OHG, Münster.

[23] Peitz, C., 2015. Systematik zur Entwicklung einer produktlebenszyklusorientierten Geschäftsmodell-Roadmap. Verlagshaus Monsenstein und Vannerdat OHG, Münster.

[24] Pynnönen, M., Hallikas, J., Ritala, P., 2012. Managing Customer-Driven Business Model Innovation. International Journal of Innovation Management 16 (4).

[25] Rose, H., 2015. Methode zur agilen Geschäftsmodell-Innovation. Dissertation, Stuttgart.

[26] Sosna, M., Trevinyo-Rodríguez, R.N., Velamuri, S.R., 2010. Business model innovation through trial-and-error learning : the Naturhouse case. International journal of strategic management (Vol. 43.), 383–407.

[27] Teece, D.J., 2010. Business Models, Business Strategy and Innovation. Long Range Planning 43, 172–194.

[28] van der Pijl, P., Lokitz, J., Solomon, L.K., 2016. Design a better business: New tools, skills, and mindset for strategy and innovation. Wiley, Hoboken, 269 pp.

[29] Weiner, N., Renner, T., Kett, H., 2010. Geschäftsmodelle im "Internet der Dienste": Trends und Entwicklungen auf dem deutschen IT-Markt. Fraunhofer-Verl., Stuttgart, 68 pp.

[30] Wirtz, B.W., Daiser, P., 2018. Business Model Innovation Processes: A Systematic Literature Review. Journal of Business Models 6 (1), 40–58.

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.

Dominik Arnold (*1991) completed a cooperative engineering education in mechanical engineering at the Bochum University of Applied Sciences and worked at the company Eickhoff Bergbautechnik GmbH. Since 2019, he has been working as a research assistant at the Chair of Production Systems (LPS) in the field of production management and is conducting research in the area of product service systems.

Christian Köhler (*1980) has been professor of industrial engineering and management as well as scientific director of the Institute of Industrial Engineering (wi institut) at Saarland University of Applied Sciences (htw saar) since 2016. Previously, Prof. Dr.-Ing. Christian Köhler worked as head of the production and engineering department for the global company Festo.

Tobias Mahl (*1992) has been working as a research assistant at the Institute of Industrial Engineering (wi institut) at Saarland University of Applied Sciences (htw saar) since 2019. He earned a bachelor's and master's degree in industrial engineering at Saarland University of Applied Sciences (htw saar). His primary research topics are business model innovation and product service systems.

Christopher Prinz (*1985) studied mechanical engineering at the Ruhr-University Bochum. After receiving his doctorate in 2018 on the topic of knowledge management in production, he was named Academic Councilor at the Chair of Production Systems (LPS). As part of the chair management, he is responsible for the strategic development of the chair and the initiation and controlling of research projects.

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.

2nd Conference on Production Systems and Logistics

Transformation Of Change Requests Into IT-Modules of ERP- and ME-Systems

Andreas Külschbach, Günther Schuh, Volker Stich
Institute for Industrial Engineering (FIR) at RWTH Aachen University, Aachen Germany

Abstract

Industrial practice shows a strong trend towards digitalization. It is not only economic crises, such as those triggered by Covid-19, that are reinforcing this trend. It is also the entrepreneurial urge to fulfill customer wishes in the best possible way and to adapt to new requirements as quickly as possible. Due to the advancing digitalization, the role of business application systems in manufacturing companies is therefore becoming increasingly important. The data processed in IT-Systems represent a great potential, especially for the evaluation of change requests in production. Through efficient change management, companies can record and process changes quickly. However, the necessary data basis to decide on existing change requests is still hardly used. Existing IT-Systems for change management coordinate the processing of change requests, but do not relate to data of operational application systems such as Enterprise-Resource-Planning. Therefore, a conceptual approach is required for the evaluation of change requests. This approach is based on an objective recording system that enables the transformation from the change description to an evaluation space. The paper presents an approach for the systematic transfer of requirement characteristics into the world of operational IT-Systems.

Keywords

Change Request; Change Management; ERP; Evaluation

1. Introduction

1.1 Initial situation

The industrial revolutions have brought many changes within production and markets. The international orientation of many companies leads to the exploitation of new markets and new customer segments. The exploitation of such markets is linked to changing competitive situations with competitors. New markets can also show enormous differences in customer requirements. In addition, the market situation can change more rapidly than ever before due to advancing digitalization [1]. The customers who consume the products and services are constantly creating new requirements. According to the Kano model, the requirements of users are continually shifting. The model clearly shows how old products and product features are considered obsolete after a certain time and are taken for granted by the customer. [2] These trends in the development of markets have accelerated significantly in recent years. Digitization promotes the exchange of information both on the customer side and on the side of the manufacturers of products and services. The customer and his requirements for a product define the market success and thus the competitiveness of a company. This is why customer centricity is more important than ever before. [3]

These developments have led to enormous pressure on product development and production cycles. Time is elementary for the development of competitive advantages. Therefore, the product development times are shortened. This is to ensure that customer requirements are fulfilled. [4] The difference between customer requirements and product functions must be kept as small as possible. The shortening of development times for new products goes hand in hand with a shortening of the utilization phase and thus a shortening of the product life cycle. [5]

1.2 Problem definition and motivation

The initial situation explained in the previous section leads to major problems for manufacturers. Highly competitive markets lead to increased pressure on companies to meet customer requirements in the best possible way. This leads to a high frequency of changes. The changes can affect the product directly or the structures that are used for manufacturing [6]. Looking at the manufacturing industry, a study from Germany attributed a significant role to change management for the 2020s [7]. One of the reasons for this is the high complexity of change management. It is a holistic approach that takes into account correlations between the areas of organization, technology, process and product [8] [9]. Moreover, this correlation must be considered in the digital representations of production. In manufacturing advancing digitization is continuously generating data streams, which are coordinated using operational application systems. It is particularly the use of data that characterizes the production and value creation processes [10]. In terms of change management, such consistent use of data is currently not evident. The systems established on the market focus on the management of change. This means that the handling of a change request is prioritized, and the systems rarely have interfaces to the operational application systems of the production and development department. Consequently, the data generated by the production system is not directly integrated into the management process of a change. The lack of a definition for a standardized change request evaluation process is one of the reasons for the insufficient integration. A comparison with change management norms and standards shows that the evaluation process is mentioned by most of them but is not specified in greater detail. The reference process of change management of the German Association of the Automotive Industry serves as an example at this point. This is explained in the associated guideline (VDA 4965) and is shown in Figure 1. It is noticeable that the assigned reference step does not contain any associated process descriptions for the evaluation process [11] [12]. This lack of integration and consideration of operational application systems in change management standards and norms is the reason for the missing integration into change management systems. This means that the people who decide whether to accept or reject a change request must rely on their expertise or search for information in the operational application systems to take it into account.

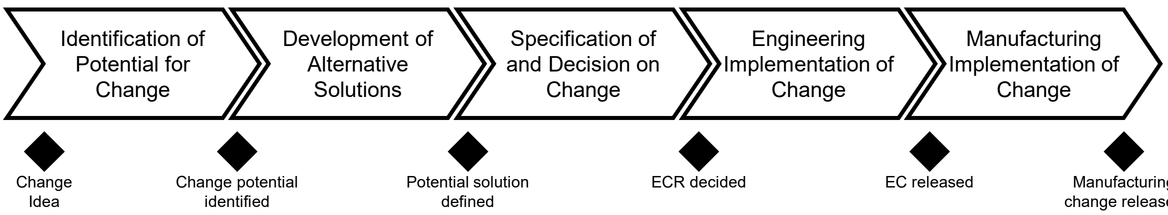


Figure 1: Engineering Change Management reference process [11]

The aim of the present approach is therefore to demonstrate a way in which change requests can be made evaluable with the help of operational application systems in use. This requires an approach that enables a connection between an existing change request and functions of the operational application systems to comply with the holistic approach in the sense of change management. For this purpose, first the objective recording of a change request is explained. It is followed by a concept for establishing a connection to the IT systems based on the recorded characteristics of a change request. For this purpose, an approach is presented that uses the Aachen-PPS-model as a reference system to link the recorded characteristics with the functions of the operational application systems. In conclusion the results are discussed and further need of research is highlighted.

2. Conception and detailing of the solution approach

The focus of the following sections is to integrate functions of operational application systems into the reference process of change management. To evaluate a change request the decision-relevant information must be established. This information can be found within an Enterprise-Resource-Planning System (ERP) and Manufacturing-Execution-System (MES) [10].

2.1 Neutral objective recording of changes

The handling of changes is usually regulated internally. In most cases, it is precisely defined according to which specifications a request must be recorded and in which order it must be processed in accordance with internal change management [8] [11]. The analysis of common standards shows a wide use of evaluation categories. In industrial practice, a combination of descriptive text in combination with defined parameters (mass, costs, cycle time, etc.) is often used for data recording. In order to standardize the recording process, companies have created templates to enable standardized recording. [13] An analog recording is the most frequently used type of recording. Change management systems increase the degree of digitization at this point by digitally representing the recording documents. In change recording, it is common to estimate effects, which is why a subjective value is the basis of the recording. This subjectivity cannot be corrected in the subsequent evaluation process. Therefore, a systematic approach is required to carry out a fact-based evaluation after the recording. For the objective recording of a change request, the approach developed by KÜLSCHBACH ET AL. is used. This uses the methodology of the characteristic schema to generate standardized description characteristics and can be seen in Figure 2 [14]. The approach is useful for executing a transformation due his standardized characteristic values which avoid textual descriptions by users. It excludes the subjective interpretation. This enables a consistent transformation.

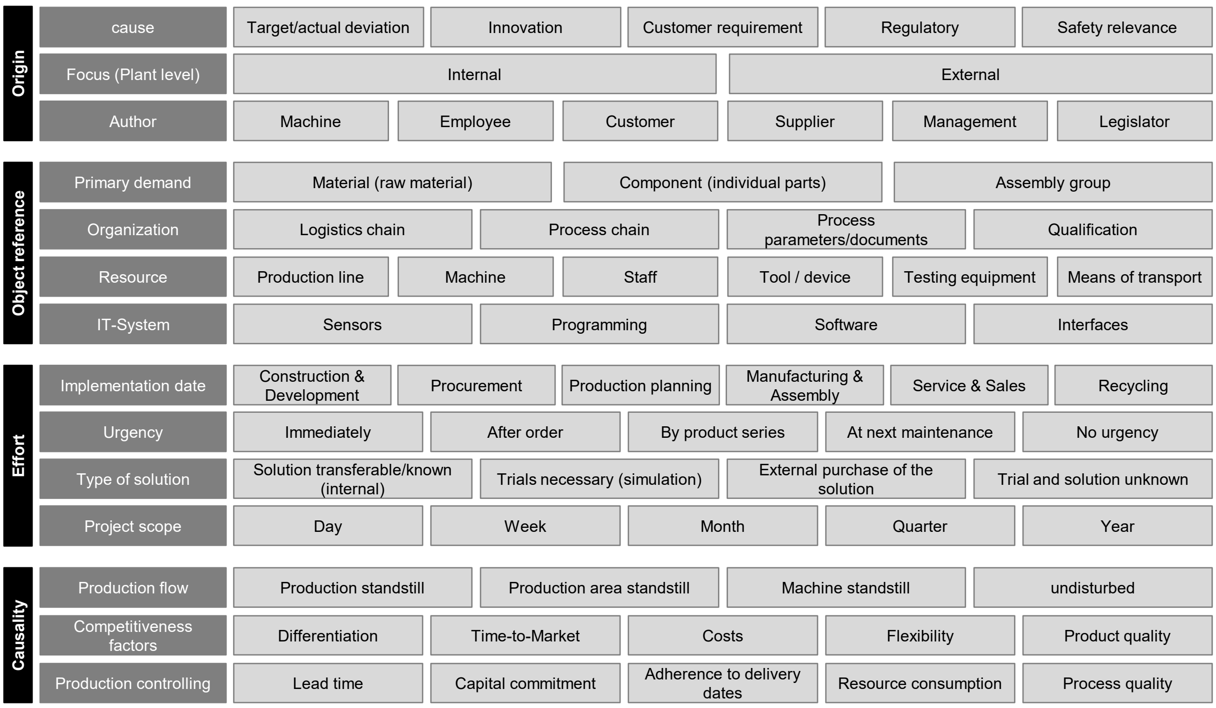


Figure 2: Characteristics schema for change request recording [14]

2.2 Transformation

The change management systems do not provide a direct link to facts or information but focus, as described in the introduction, on the pure management and coordination of change requests. This creates a gap in the information flow, which makes the evaluation even more difficult and is not considered user-friendly. To be able to perform a transformation from the level of the characteristic schema according to KÜLSCHBACH ET AL. a process-related reference is required. Operational application systems are characterized by their support function within order processing. Therefore, the processes of order processing can be understood as a connecting element between the characteristic attributes of the characteristic schema and the functions of the operational application systems. The elaborations of SCHUH ET AL. offer a suitable framework for this, as the tasks of production planning and control are defined as a descriptive reference model [15]. The analysis of the publications of SCHUH ET AL. shows a deficit in the integration of IT-Systems. The wide distribution and use of the approach therefore shows the need for research [16] [17] [18].

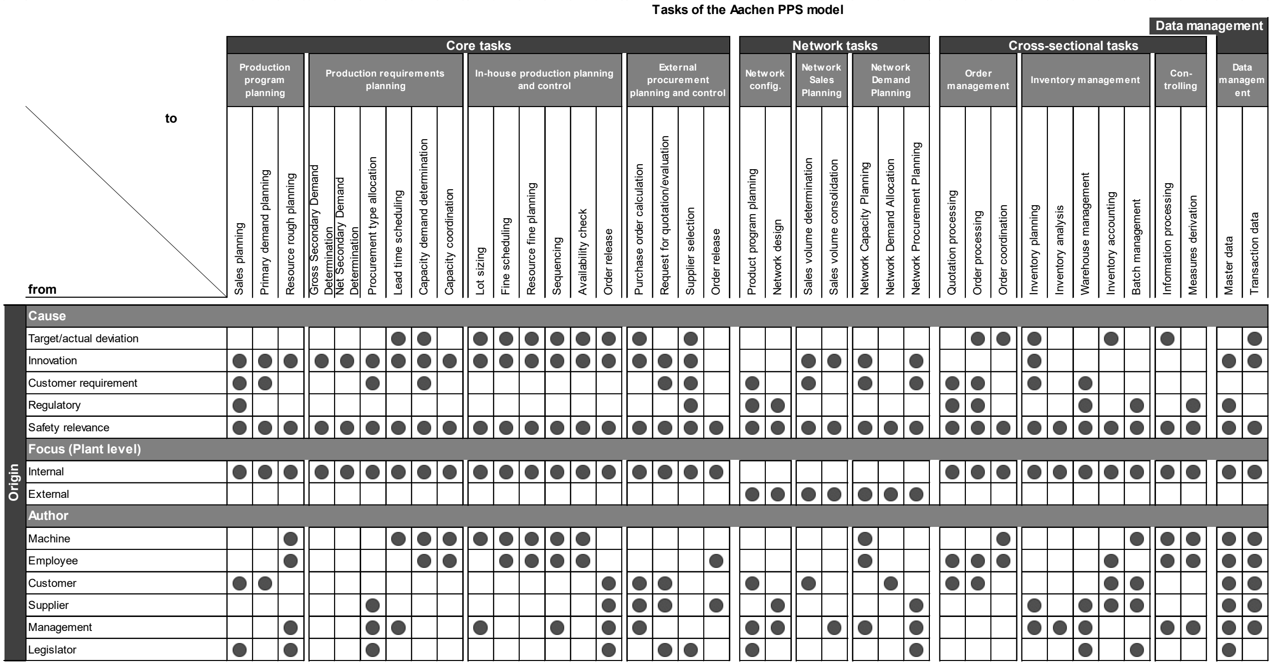


Figure 3:Transformation matrix for characteristics schema to PPS tasks – domain *Origin*

For connecting IT-modules to the tasks of production planning and control a two-stage procedure was selected for the transformation. This is based on the TRIZ methodology according to ALTSHULLER, which is used to solve complex tasks. A concrete problem is transformed into a solution space and then concretized. This is necessary because the direct solution path from specific problem to specific solution is too complex. [19] This is also valid for a transformation from the characteristic schema to IT-modules and functions. Therefore, a generally valid reference point for the assignment of characteristic attributes to IT-modules and functions is needed. This reference is given by the tasks of the Aachener-PPS-model according to SCHUH ET AL [15]. Based on these elaborations, a logical derivation of the assignments was developed in the first step of the transformation. These correlations between characteristic attributes of the schema from KÜLSCHBACH ET AL and tasks from the Aachener-PPS-model were made based on an analysis of content description. In addition, the correlations were validated in discussions with experts from industrial practice. These experts were representatives from different ERP providers like Asseco Solutions; Comarch, SAP® and proALPHA. Figure 3 shows the correlations between the characteristic attributes of the domain *Origin* (represented within rows) and the tasks of production planning and control (represented by columns). As an example, an excerpt of the correlations of the characteristic attribute *Customer* from the *Author* domain will be explained. The customer as the source of a change has a decisive impact on *sales planning (Core task: Production program planning)* since customer requirements are decisive for this planning. From the sales plan, internal production demands, based on end products, can be identified. The characteristic attribute therefore also correlates with the tasks of *product program planning (Network task: Network config.)* and *primary demand planning (Core task: Production program planning)*. If a change is initiated by the customer, his sales orders may be affected, which in return may have an impact on the *order release* task of the in-house production planning and control. Through the correlations between characteristic schema and Aachener-PPS-model tasks, the specific problem was transferred to the solution space according to TRIZ methodology. Based on these correlations, the user can understand how the characterized change request affects certain production tasks. After that, a connection to IT systems must be established to provide the user with additional data and information. This allows effects of the associated change to be evaluated on a data basis. Therefore, the second step of the transformation, the correlations between tasks from the Aachener-PPS-model and IT-modules and functions were analyzed. For this purpose, the current performance range of ERP and MES were analyzed. The range of performance was defined by the means of requirement specification sheet, which are used for the selection of ERP and ME systems. The requirement specification sheets for ERP and MES were provided by Trovarit AG. The Trovarit AG supports national and international companies in the selection of ERP and ME systems. It has a platform on which over 700 ERP providers and over 350 MES providers are listed to participate in tenders. The analyzed requirement specification sheets are used for these tenders.

As a result of the analysis, connections are made between the tasks of the Aachener-PPS-model and modules of the IT-Systems. At this point, correlations are to be clarified based on the task *In-house production planning and control*. This core task of the Aachener-PPS-model is characterized by controlling and planning activities. These includes planning functions such as *detailed resource planning* and *sequencing of production orders*. These planning functions are described in the *production planning* IT-module which is why this IT-module is clearly related to the core task *In-house production planning and control*. Furthermore, a comparison is made with the current stock levels and stock status as part of the material availability check. This clearly shows that *material and warehouse management* as an IT-module is affected to the *In-house production planning and control* core task. As a result of the analysis, correlations to modules of *production planning*, *material master data management*, *material and warehouse management*, as well as *shipping* can be determined. The complete result of the analysis for all tasks of the Aachener-PPS-model is summarized in Figure 4.

Tasks of Aachener PPS-Model	Correlating IT system modules
Core tasks	
Production program planning	Sales; Production planning
Production demand planning	Disposition; Purchasing & Procurement; Material master data management; Production planning
In-house production planning and control	Material master data management; Control level; Shipping; Materials & warehouse management; Production planning
External procurement planning and control	Purchasing & Procurement; Materials & warehouse management
Network tasks	
Network configuration	Production planning; Disposition; Purchasing & Procurement; Sales
Network Sales Planning	Sales
Network Demand Planning	Purchasing & Procurement; Production planning; Disposition
Cross-sectional tasks	
Order management	Sales; Service & Repair; Production planning; Materials & warehouse management
Inventory management	Disposition; Control level; Materials & warehouse management; Material master data management; Production control
Controlling	Control level; Materials & warehouse management; Material master data management; Production control
Data management	Total scope

Figure 4: Assignment of IT-modules to production planning and control tasks

Based on this allocation, the detailing of the identified IT-modules can be carried out for all tasks of the Aachener-PPS-model. For this purpose, the requirements specification sheet of an ERP and an ME system were analyzed in detail for every IT-module [20] [21]. To show the correlations, the core tasks were summarized in rows and the IT-modules with their functions are summarized in columns as shown in Figure 5. In the following, the scope of the core task *Production demand planning* will be explained in excerpts. The aim of *Production demand planning* is the realization of the production program. The determination of demand can be carried out deterministically with the help of bill of materials (BOM) explosion. These are stored in the material master data, which is why there is a correlation to *Material master management* and *BOM management* within the IT function module *Material master data management*. For the *Disposition* module, a correlation to *Material disposition* can be proven since this is decisive for the determination of heuristic requirements. Since no allocation with stocks within the *Gross secondary demand determination* takes place, no correlations can be proven in the module *Purchase and Procurement*. The *Net secondary demand determination* determines the necessary dependent demands to be procured. This takes place by consideration of different stock types, which are planned either on date or summarized within a period. This justifies the correlation to *Production program planning* and *Production planning*. Since material master data are also used, the correlations of the *Gross secondary demand determination* for the individual functions *Material master management* as well as *Material planning* are also valid. In addition, the result is the material quantity to be procured. Therefore, a correlation must exist in the context of *Order quantity determination* and *Purchase requisitions*. The subtask *Procurement type allocation* must determine for the materials to be procured, according to which procurement type is acted. This is stored in the material master data, which is why there is a correlation with the function *Material master management*. The strategic *Make-or-buy decision* is a core component of *Production demand planning*, which is why the associated task correlates. *Lead time scheduling* structures the production process from a time perspective. Compared to detailed *In-house production planning and control*, rough key dates are targeted. For this, an unlimited capacity of resources is assumed. *Resource management* is therefore directly affected. The reference to production orders is established in the context of *Production planning*, which additionally enables the possibility of a first rough *Production simulation*. For the subtask *Capacity demand determination*, the required available capacity is determined. The processing time depends on the unit lead time and the number of units. Therefore, a correlation to *Lot sizing* can be identified. The entirety of the correlations is shown in Figure 5.

Functional modules of operational application systems for production

to / from	Material master data management				Disposition			Purchasing and Procurement								Production planning					
	Material master management	BOM management	Connection PDM/CAD	Variant Management	Make-or-Buy-Decision	Material disposition	Kanban	Supplier Management	Purchase requisitions	Order quantity determination	Purchase order processing	Purchase order monitoring	External production	Invoice verification	Supplier frame orders	Resource management	Work preparation	Lot sizing	Production program planning	Production planning	Production simulation
Core tasks – Production demand planning																					
Gross Secondary Demand Determ.	●	●				●													●		
Net Secondary Demand Determ.	●					●			●	●									●	●	
Procurement type allocation	●				●			●			●		●							●	
Lead time scheduling																●				●	●
Capacity demand determination																●		●		●	
Capacity coordination																●	●			●	●

Figure 5: Correlation between Production demand planning and IT-Modules

3. Outlook and need for further research

The results of the study are two transformation matrices which enable a transfer to modules and functions of the operational application systems ERP and MES. The approach generates a standardized procedure for transforming and evaluating change requests. Users can link the change request with data and information from operational application systems. The approach presented in this paper enables users to understand the linkage between tasks of the production system and specific IT functions. The first transformation matrix is used to transform an objectively recorded change request. The use of the Aachener-PPS-model guarantees a high degree of transferability to manufacturing companies. The second transformation enables the user to identify affected IT functions. Discussions with providers of ERP systems and experts from industrial practice show that the use of standardized requirement specification sheets enables a high degree of transferability to industrial practice. The approach presented in this paper can therefore be applied regardless of the IT-Systems used. This allows an integration into the change management processes of industrial companies.

Further research must be conducted to define the preparation for decision support. Regarding the usability of the presented approach, special attention must be paid to the management suitability. It is necessary to describe the impact effects of a change request in a standardized way. Only an objective presentation of the evaluation result guarantees the reliability of a following evaluation process. In conclusion, the combination of objective recording, transformation and evaluation allows a comprehensive and objective evaluation of change requests. The approach presented in this paper contributes to the further development of production systems that can adapt more quickly to changing customer requirements.

References

[1] Lauer (2021) Change management: Fundamentals and success factors. Springer, Berlin, [Heidelberg]

[2] Coleman (2015) The customer-driven organization: Employing the Kano Model. A productivity press book. CRC Press, Boca Raton

[3] Plehn 2017 A Method for Analyzing the Impact of Changes and their Propagation in Manufacturing Systems. Dissertation, Technische Universität München

[4] Schuh et al. (2017) Enabling Agility in Product Development through an Adaptive Engineering Change Management. Procedia CIRP 63: 342–347. doi: 10.1016/j.procir.2017.03.106

[5] Vernon 1966 International Investment and International Trade in the Product Cycle. The Quarterly Journal of Economics 80(2): 190–207

[6] Smith (2017) Reliability, maintainability, and risk: Practical methods for engineers, Ninth edition. Butterworth-Heinemann, Kidlington, Oxford, United Kingdom, Cambridge, MA

[7] Koch et al. (2015) Änderungsmanagement in der Produktion, München

[8] Becerril et al. 2016 Estimating the effects of Engineering Changes in early stage product development. 18th INTERNATIONAL DEPENDENCY AND STRUCTURE MODELING 2016(18): 125–135

[9] Bauer et al. 2020 Modular change impact analysis in factory systems. Prod. Eng. Res. Devel. 14(4): 445–456

[10] Achatzi et al. 2020 Application and Information Integration. In: Vajna S (ed) Integrated Design Engineering. Springer International Publishing, Cham, pp 549–585

[11] Verband der Automobilindustrie (2010) ECM Recommendation Part 0 (ECM): VDA 4965 – Part 0

[12] Verband der Automobilindustrie (2010) ECM Recommendation Part 1 (ECR): VDA 4965 – Part 1

[13] Macke et al. 2016 Standardization of Smart Manufacturing Change Management. In: Goh YM, Case K (eds) Advances in manufacturing technology XXX: Proceedings of the 14th International Conference on Manufacturing Research, incorporating the 31st National Conference on Manufacturing Research, September 6 - 8, 2016, Loughborough University, UK. IOS Press, Amsterdam, pp 541–546

[14] Külschbach et al. 2020 Charakterisierung von Änderungsanfragen in Produktionssystemen. ZWF 115(11): 765–768

[15] Schuh et al. (2012) Produktionsplanung und -steuerung 1. Springer Berlin Heidelberg, Berlin, Heidelberg

[16] Lödding (2016) Verfahren der Fertigungssteuerung: Grundlagen, Beschreibung, Konfiguration, 3. Auflage. VDI-Buch. Springer Vieweg, Berlin, Heidelberg

[17] Wiendahl et al. (2019) Betriebsorganisation für Ingenieure: Mit 279 Abbildungen, 9., vollständig überarbeitete Auflage

[18] Westkämper et al. 2018 Fertigungs- und Fabrikbetrieb. In: Grote K-H, Bender B, Göhlich D (eds) Dubbel: Taschenbuch für den Maschinenbau. Springer Berlin Heidelberg, Berlin, Heidelberg, S128–S162

[19] Springer International Publishing TRIZ. Theory of inventive problem solving: Level 1. TRIZ from A to Z

[20] Bach et al. (2019) Marktspiegel Business Software ERP/PPS 2019/2020: Anbieter - Systeme - Projekte, 10. Auflage, revidierte Ausgabe. Trovarit, Aachen

[21] Wiendahl et al. (2019) Marktspiegel Business Software - MES - Fertigungssteuerung 2019/2020. Trovarit

Biography

Andreas Külschbach M.Sc. (*1991) has been working as a project engineer at FIR at RWTH Aachen University since 2017. In his current position as head of Production Planning as part of the Production Management Division, he supports companies in various industries in the design and implementation of efficient production and logistics systems. He 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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Requirements For Incentive Mechanisms In Industrial Data Ecosystems

Joshua Gelhaar[1], Jan Ruben Both[1], Boris Otto[1]
[1]Fraunhofer Institute for Software and Systems Engineering ISST, Dortmund, Germany

Abstract

In the increasingly interconnected business world, economic value is less and less created by one company alone but rather through the combination and enrichment of data by various actors in so-called data ecosystems. The research field around data ecosystems is, however, still in its infancy. In particular, the lack of knowledge about the actual benefits of inter-organisational data sharing is seen as one of the main obstacles why companies are currently not motivated to engage in data ecosystems. This is especially evident in traditional sectors, such as production or logistics, where data is still shared comparatively rarely. However, there is also consensus in these sectors that cross-company data-driven services, such as collaborative condition monitoring, can generate major value for all actors involved. One reason for this discrepancy is that it is often not clear which incentives exist for data providers and how they can generate added value from offering their data to other actors in an ecosystem. Fair and appropriate incentive and revenue sharing mechanisms are needed to ensure reliable cooperation and sustainable ecosystem development. To address this research gap and contribute to a deeper understanding, we conduct a literature review and identify requirements for incentive mechanisms in industrial data ecosystems. The results show, among other things, that technical requirements, such as enabling data usage control, as well as economic aspects, for instance, the fair monetary valuation of data, play an important role in incentive mechanisms in industrial data ecosystems. Understanding these requirements can help practitioners to better comprehend the incentive mechanisms of the ecosystems in which their organisations participate and can ultimately help to create new data-driven products and services.

Keywords

Requirements; incentives mechanisms; data sharing; industrial data ecosystems

1. Introduction

The steady spread of information and communication technologies leads to the fact that data is increasingly a driver of change and economic growth [1]. Whereas in the past data was primarily used to improve internal processes, today it more and more serves as a strategic resource that forms the basis for the development of data-driven innovations and business models [2]. At the same time, however, data generation and data processing are neither the core competence nor the core business of most companies. In addition, it is increasingly no longer just the company's own data that is of interest [3]. On the one hand, large amounts of data are needed for meaningful analysis purposes, which is further reinforced by the trend towards artificial intelligence. On the other hand, the required data is often not generated within the respective company itself. As a result, data-driven innovation and economic value creation are less likely to be created by individual organisations or in traditional value chains. Instead, data-driven value creation takes more and more place in cross-industry, socio-technical networks - so-called data ecosystems [4]. This development implies that participation in data ecosystems is becoming increasingly likely and relevant for companies. Some authors even claim that ecosystem engagement is an urgent necessity rather than a choice for companies [5]. However, many companies are still reluctant to share their data with others and are therefore unable to take advantage of the potential and benefits that arise from participating in data ecosystems [6]. In particular in traditional sectors, such as production and logistics, data is still comparatively rarely shared across companies [7]. One main reason for this is seen in the fact that it is often not clear what incentives exist and ultimately how data providers can benefit when they offer their data to other actors in an ecosystem [8]. For this reason, a functioning and sustainable ecosystem requires fair and appropriate incentive mechanisms which motivate actors to participate in a data ecosystem [9]. To address this research gap and contribute to a better understanding about the structure of incentive mechanisms in industrial data ecosystems, we aim to answer the following research question in this paper:

Research question: *What are requirements for incentive mechanisms in industrial data ecosystems?*

The remainder of the paper is structured as follows: First, we give an overview of the theoretical concepts of industrial data ecosystems and incentive mechanisms. Section 3 describes our structured literature review and analysis process. Afterwards, we outline the identified requirements for incentive mechanisms in industrial data ecosystems. Finally, we conclude the paper with a discussion of the results and provide an outlook on future research topics.

2. State of the art

2.1 Use case example: Collaborative condition monitoring

Condition monitoring is the process of regular or permanent monitoring of a machine condition by measuring and analysing physical parameters such as vibration or temperature. The goal is to analyse the data from the machine sensors to detect behavioural patterns that may indicate a developing fault in the machine. This data is classically shared bilaterally, e.g. exclusively between the machine manufacturer and the machine operator. The idea of collaborative condition monitoring (CCM) is that data is not only shared bilaterally but multilaterally between all actors in an ecosystem [10]. This increases the amount of data available which in turn improves the results of the data analysis. Figure 1 gives an overview of a CCM use case. The component supplier shares lifetime or reliability relevant data about the machine components produced by him. In return, he gains access to operating data for his components and other relevant associated machine data. This enables the optimization of his components or new services, such as proactive spare parts management. Based on historical data from many machines in a wide variety of environments, the machine manufacturer can use AI methods to recognize how, for example, the availability and tolerances of the machines change in production. With this knowledge, the manufacturer can proactively contact the machine operator with a maintenance offer and generate increased customer satisfaction. For this, however, the machine manufacturer needs the operating and environmental data of the machine operator. The machine operator provides this data and benefits from the predictive maintenance service of the machine manufacturer with increased machine availability. However, the potential of CCM can only be realized in practice if data is shared by as many actors as possible. This also includes actors who may not generate any directly apparent added value by sharing their data. For this reason, incentive mechanisms are needed that address all actors and motivate them to participate in collaborative data sharing [10].

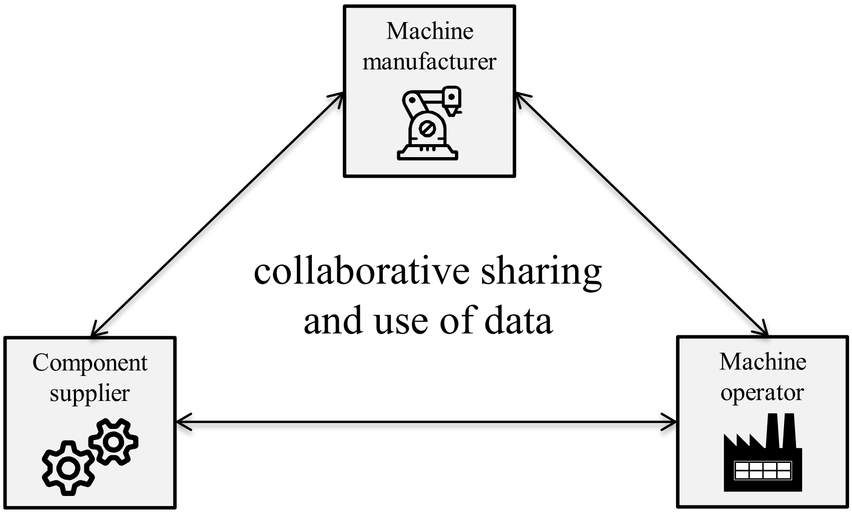


Figure 1: Collaborative condition monitoring use case overview [10]

2.2 Industrial data ecosystems

Data is often seen as the raw format that becomes information through processing. This information is then combined and interpreted to generate knowledge that can be used for decision-making processes or new business models [11]. This value creation process is also known as the data value chain [12]. A data value chain generally consists of the three phases: data collection, data interpretation, and data exploitation [8]. Data collection includes the generation and acquisition of data by e.g. sensors [13]. Subsequently, data analysis methods such as machine learning are used to process raw data into useful information. Lastly, in the exploitation phase, the information is integrated into business activities and translated into business value such as cost reduction [12]. Due to the development that data-driven services and products are increasingly based on the combination of multiple data sources from various actors, the different phases of the data value chain are often performed by different actors instead of one resulting in new forms of cross-company collaboration [14]. These forms of collaboration are also referred to as data ecosystems which, consequently, focus on the cross-actor generation, processing, and use of data with the aim of creating added value for all actors involved [15,7]. In this context, data ecosystems can generally emerge in the three domains scientific, government, and industry [16]. Since data ecosystems consisting of industrial companies in particular have not been well studied, we focus on industrial data ecosystems in this paper [17]. Data ecosystems are characterized by, among other things, complex interdependencies among their participants leading them in some cases to work cooperatively and competitively at the same time, which is also known as coopetition [18]. Based on the data at the focus of an ecosystem, different actors have varying relationships to it. This leads to different roles with a variety of functions that can be taken on [16]. At a minimum, there are the following three roles: Data providers who collect and provide data, analytics service providers who analyse the data, and data consumers who perform data exploitation activities [8].

2.3 Incentive mechanisms in industrial data ecosystems

Incentive mechanisms are studied in many different fields when it comes to encouraging people or organisations to do something in exchange for a reward. Incentive mechanisms in industrial data ecosystems are important for the following reasons: As described above, data-driven products and services are increasingly based on the combination of different data which originate from both internal and external data sources [19]. However, many companies are reluctant to share their data with others as the data may contain confidential and valuable information and sharing it could consequently strengthen competing companies [20]. Second, it is difficult to valuate what a data provider should get back for its data, e.g. in the form of money, as there are still no standardized methods for the monetary valuation of data both in research and practice [7,21]. Finally, all activities in the data value creation process incur costs, e.g. for maintaining the data and its quality along the data lifecycle [8,22]. Similarly, making data available for sharing creates costs in terms of effort and time for the data providers [23]. However, revenues are only generated in the data exploitation phase [8]. For these reasons, the expenses for all other activities that can be performed by different actors must be compensated by the revenues from the data exploitation activity in such a way that a win-win situation is achieved for all actors involved [24,8]. This compensation can be done in different ways such as a direct payment or indirectly through, for example, a service [8]. Incentive mechanisms should ultimately ensure that each actor is appropriately rewarded for sharing their data, to promote fair and sustainable collaboration in the ecosystem [25].

3. Research method

The objective of this study is to identify requirements for incentive mechanisms for data sharing in industrial data ecosystems based on findings from the relevant literature. We, therefore, performed a structured literature review following the approach by [26] and the guidelines by [27]. 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 great results for our field of interest, indexing the most relevant journals and conference proceedings [28]. We performed an initial search using the keywords "data ecosystem" AND "data sharing" AND "incentive". This, however, resulted in no hits. To overcome this shortcoming, we decided to expand and simplify the keywords. Therefore, we used "data sharing" OR "data exchange" AND "incentive" as our search query. We added the term "data exchange" because it is used synonymously with the term "data sharing" by some authors in the literature [29]. The results were limited to English language literature and peer-reviewed only. As a first set, this resulted in a number of 344 papers. Within this initial set of papers, we examined the titles, abstracts, and keywords in terms of relevance to our research question. We eliminated papers, for example, dealing with incentive mechanisms for sharing research data as conditions and incentives within these research data ecosystems are different from those in industrial data ecosystems [16]. This filtering process resulted in 17 relevant papers. We then performed a forward and backward search as suggested by [30] which resulted in 10 additional relevant articles. In addition, we added the 15 articles on industrial data ecosystems from a recent systematic review of the data ecosystem literature in which the authors selected and reviewed articles based on additional search terms in other prominent bibliographic databases [16]. This resulted in a total of 42 articles as the basis for the literature analysis process (see Figure 2).

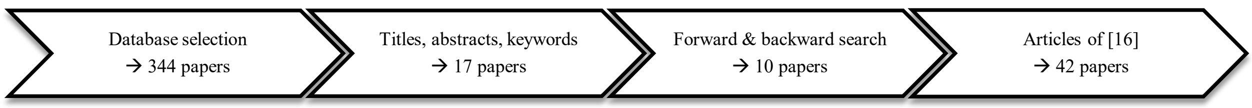


Figure 2: The structured literature review process

For the literature analysis, we followed an explorative approach to identify relevant similarities and interesting facts to answer our research questions [30]. The rationale for choosing this exploratory approach is the lack of theories on incentive mechanisms in industrial data ecosystems. To code the literature, we followed the grounded theory coding process [31]. Accordingly, open coding was first used to label the literature with categories that summarize the relevant content. Subsequently, relationships between the categories were identified by axial coding. Last, selective coding was applied to aggregate the identified categories into more general dimensions. In summary, we obtained five aggregated dimensions that represent, respectively, requirements for incentive mechanisms in industrial data ecosystems. These five requirements are described individually in detail below.

4. Results

4.1 Requirement 1: Ensure quid pro quo

One of the main requirements of incentive mechanisms mentioned in the literature is to ensure that each actor has to benefit through its participation in the data ecosystem [20]. Otherwise, it would not be rational for a company to participate in the ecosystem if the participation would lead to a worse company state [8]. Here, quid pro quo means that every actor who puts something into the data ecosystem, in the form of data, money, or any other effort, must also get something back. Particularly for data providers, it must be unambiguously and transparently clear how they can profit by sharing their data [24]. This requires that data providers must be adequately compensated for their data, for the costs incurred in collecting and storing the data, and for the risk they take in sharing it [8]. In general, therefore, it is necessary to consider within the ecosystem how the revenues, which are usually generated only when data are exploited, are distributed fairly [25]. For this it is required that the shared data is valuated in monetary terms, which is still a great challenge due to the special characteristics of data [7]. That is why some authors propose models that take into account proportionally how much contribution the shared data has on the revenue or lead to an improvement of the data-based service [13,8]. However, these models also have the challenge of calculating this proportionate added value of the total value. This is especially difficult when multiple data sets from different data providers are combined and analysed by one data consumer [8,25]. Nevertheless, a distribution of profits or the compensation of efforts does not necessarily have to be monetary for all actors. Instead, an actor may receive other forms of payment in return for their data, such as data or a service [24]. For these reasons, it is important to identify and quantify what added value each actor can receive through their participation in the data ecosystem. In this regard, there are often cases in which one's own added value is not directly recognizable for some actors [24]. If this is the case, the other ecosystem actors should transparently present and demonstrate this possible added value, e.g. by showing possible improved economic key performance indicators [13]. In literature and practice, there are increasing ideas to solve these challenges of profit distribution with blockchain-based tokens as these can be exchanged digitally more easily and securely than money, for example [32].

4.2 Requirement 2: Improve data quality

A second requirement identified in the literature is the incentive to share data with a high quality in the ecosystem. Since data is the basis for data ecosystems and the added value they create, it is important that the data in the ecosystem has a high quality, e.g. is correct and consistent [33]. However, data quality is context-dependent, which is why data providers often do not know what data quality is required by data consumers [34]. For this reason, on the one hand, there must be sufficient opportunities in the ecosystem for communication and feedback between data providers and the other actors, such as the analytics service providers, so that data quality can be continuously improved [16]. In some situations, it may also be useful to share poor quality data and let other actors or a crowd use and improve it [35]. On the other hand, all data-providing actors must be incentivized to provide high-quality data to the ecosystem. Higher quality data generally serves as a better basis for service providers' data-based analytics, who consequently can generate higher revenues with higher-quality data [13,8]. For this reason, among others, higher data quality is associated with higher monetary data value [22]. Based on this, it seems reasonable and necessary to incentivize high data quality within the data ecosystem, e.g. through monetary incentives. Compensating the effort for high data quality can also lead to network effects since a data set with a high quality has a higher chance to be used in the ecosystem and could also be relevant for several other actors [8].

4.3 Requirement 3: Establish trust

Another requirement for incentive mechanisms frequently mentioned in the literature is the establishment of trust. Trust is central to motivate actors to participate, e.g. by sharing their data in a data ecosystem [13,24].

On the one hand, this can be trust towards other actors in the ecosystem and for which purposes they will further use the shared data [36]. This is also often associated with trust in the technical infrastructure that is used for data sharing and the requirements it can fulfil, for example, with regard to data security and sovereignty. On the other hand, this can be trust in the actual shared data and its quality, as poor data quality can lead to significant damage in e.g. business processes [24]. These different types of trust in the data ecosystem can be established through a number of ways and methods. Firstly, these can be technical measures that regulate data security, data access, and data usage rights, for example, or track where data comes from and goes to [20]. This kind of trust can be established through the use of trust-building technologies, such as distributed architectures and ledger technologies, or through program verification and certification [9]. Secondly, trust can also be established through legal measures, such as contracts, or organisational and governance measures, such as joint agreements and rules for the ecosystem [37]. Another way to build trust in data ecosystems is to establish a trustee, such as a data fiduciary, who can serve as an independent intermediary between the ecosystem actors such as data providers and data consumers [9,38]. This concept is based, among other things, on a good reputation that the actors have towards this third party. In general, reputation is mentioned by some authors as another measure to build trust. For example, [39] introduce a blockchain-based infrastructure that allows data providers to be rated, which in turn can lead to data providers with better ratings being trusted more and vice versa.

4.4 Requirement 4: Foster sustainable ecosystem building

The data ecosystem concept is based on the idea of companies combining their individual offerings and data into an integrated solution that enables actors to create value through joint efforts [7]. Consequently, each actor hopes that the common goal is greater than the sum of its individual parts, and thus that they are better off by participating in the ecosystem than if they were alone [8,40]. For a functioning and sustainable data ecosystem, it must therefore be ensured that, on the one hand, each actor has an individual business model and a business strategy within the ecosystem and, on the other hand, that the ecosystem as a whole has common goals and value propositions [33,6]. Conversely, individual interests of the various actors and the complex relationships between them can lead to conflicting goals in the ecosystem [16]. For example, maximizing revenue for a data consumer may conflict with maximizing welfare for the entire ecosystem [8]. In addition, the ecosystem actors need to cooperate in some areas that do not directly add value or generate revenue, such as the joint development of standards for data models and interfaces [8,20]. Ultimately, a successful and sustainable data ecosystem must ensure that its members have a shared understanding of the ecosystem's operations and goals as well as that each actor is satisfied with its position in the ecosystem [41]. Sustainability in this context also means that the data ecosystem considers how it wants to develop further in the future and how, for example, new actors are accepted into the ecosystem. Depending on the structure and objective of the ecosystem, different decisions can be made in this regard [37].

4.5 Requirement 5: Avoid free-riders

Another category identified in the literature deals with hitchhiking and free-riders. Analogous to the sharing of other material or immaterial goods, the problem of free riding can also be observed in the sharing of data. This can be explained, on the one hand, by the fact that data sharing can result in the loss of exclusive control over the data, and opportunism on the part of competitors can inflict great losses on the sharing parties [42]. On the other hand, analogous to other forms of cooperation between companies, spillover effects can also be expected in data ecosystems [43]. These two circumstances can reduce the willingness of companies to share their data with others and instead encourage free-riding. For these reasons, incentive mechanisms must address and prevent excessive selfish behaviour by actors and promote truthfulness [33,8]. This is particularly relevant for alliance-driven and emerging data ecosystems, as these ecosystems can only function if several actors make contributions in the form of investments, for example [9,20]. In addition, the ecosystem community should consider how to deal with free-riders or even harmful behaviour and under what circumstances actors can be excluded from the ecosystem.

5. Conclusion

Using a structured literature review, we identified requirements for incentive mechanisms in industrial data ecosystems in this paper. The analysis of the existing literature has shown that some papers have already dealt with incentive mechanisms in data ecosystems. However, authors often focus only on single aspects, such as technical matters or specific domains. To the best of our knowledge, there is no previous work that provides a comprehensive overview of these different approaches in the form of requirements as it has been done in this paper.

From our results, we can derive several implications for theory and practice. In terms of **scientific contributions**, our work firstly contributes to a deeper general understanding of the emerging and still unexplored research field around industrial data ecosystems. In addition, the systematic description of the requirements aims to expand the existing body of knowledge about incentive mechanisms in data ecosystems and to contribute to the specification of a common understanding of these complex issues. In addition, our results can help in the development of incentive mechanisms and thus be the basis for methods for the sustainable building of data ecosystems that have not yet been explored in the literature [16].

Furthermore, the results of our paper provide multiple **contributions for practitioners**. First, the results can be used by organisations to understand the incentive mechanisms in the data ecosystems in which they are already involved. A better understanding of these issues could help practitioners shape the incentive mechanisms, and thus the data ecosystem, to their advantage and ultimately generate greater value from them. Second, the results of this study can help organisations and communities to build and design data ecosystems along with their incentive mechanisms, with the goal of realizing the benefits of sharing data across organisations [9].

The results of our study are, naturally, subject to certain **limitations** which should be taken into account when interpreting them. First, the identified requirements are based on an analysis of the scientific literature on industrial data ecosystems by the author team making the data search and analysis itself subject to interpretation. Consequently, other researchers might derive different results depending on their individual influences, preferences, and predilections. In addition, consideration of practice-based insights, such as those gained in case studies, could validate or extend the paper findings. Second, it should be noted that due to the constant technological progress and the still small number of studies, the concepts and understanding around industrial data ecosystems and incentive mechanisms are constantly evolving [15]. Lastly, the lack of a clear understanding and commonly accepted definition of data ecosystems makes it difficult to distinguish between related ecosystem concepts, e.g. platform ecosystems and other related forms of collaboration, such as corporate alliances or business networks [16].

However, the above-mentioned limitations indicate opportunities for **future research** topics. A possible next step would be to analyse how the requirements described can be implemented in practice. For this purpose, some of the requirements described should be examined in greater depth and their implementation options analysed. For example, it could be investigated in more detail how trust can be established in practice via technologies and how these can be implemented, or how trust can be measured in an ecosystem [9]. In general, building on the identified requirements and following related research topics, we see the development of design principles that support practitioners in the systematic implementation of incentive mechanisms for industrial data ecosystems as a next useful research avenue.

References

[1] Zolnowski, A., Christiansen, T., Gudat, J., 2016. Business Model Transformation Patterns of Data-Driven Innovations, in: Proceedings of the 24th European Conference on Information Systems.

[2] Davenport, T.H., 2013. Analytics 3.0. Harvard Business Review 91 (12), 64–72.

[3] Lis, D., Otto, B., 2020. Data Governance in Data Ecosystems - Insights from Organizations, in: AMCIS 2020 Proceedings.

[4] Oliveira, M.I.S., Lóscio, B.F., 2018. What is a data ecosystem?, in: Proceedings of the 19th Annual International Conference on Digital Government Research, Delft, Netherlands.

[5] Thomas, L.D.W., Autio, E., 2015. The processes of ecosystem emergence. AMPROC.

[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] 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.

[9] 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.

[10] Bundesministerium für Wirtschaft und Energie (BMWi), 2020. Kollaborative datenbasierte Geschäftsmodelle: Collaborative Condition Monitoring – Wie durch unternehmensübergreifende Kollaboration Mehrwert generiert werden kann.

[11] Yates-Mercer, P., Bawden, D., 2002. Managing the paradox: the valuation of knowledge and knowledge management. Journal of Information Science 28 (1), 19–29.

[12] 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.

[13] Azkan, C., Iggena, L., Möller, F., Otto, B., 2021. Towards Design Principles for Data-Driven Services in Industrial Environments, in: Proceedings of the 54th Hawaii International Conference on System Sciences, Hawaii, USA, pp. 1789–1798.

[14] Attard, J., Orlandi, F., Auer, S., 2017. Exploiting the Value of Data through Data Value Networks, in: Proceedings of the 10th International Conference on Theory and Practice of Electronic Governance - ICEGOV '17, pp. 475–484.

[15] Gelhaar, J., Groß, T., Otto, B., 2021. A Taxonomy for Data Ecosystems, in: Proceedings of the 54th Hawaii International Conference on System Sciences, Hawaii, USA.

[16] Oliveira, M.I.S., Barros Lima, G.d.F., Lóscio, B.F., 2019. Investigations into Data Ecosystems: a systematic mapping study. Survey Paper. Knowledge and Information Systems.

[17] Gröger, C., 2021. There Is No AI Without Data: Industry Experiences on the Data Challenges of AI and Call for a Data Ecosystem for Industrial Enterprises. Communications of the ACM, to appear.

[18] Moore, J.F., 1993. Predators and Prey: A New Ecology of Competition. Harvard Business Review 71 (3), 75–86.

[19] 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).

[20] Gelhaar, J., Otto, B., 2020. Challenges in the Emergence of Data Ecosystems, in: Proceedings of the 24th Pacific Asia Conference on Information Systems (PACIS).

[21] 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.

[22] Otto, B., 2015. Quality and Value of the Data Resource in Large Enterprises. Information Systems Management 32 (3), 234–251.

[23] Prüfer, J., 2020. Competition policy and data sharing on data-driven markets: Steps towards legal implementation. Friedrich-Ebert-Stiftung, Bonn.

[24] 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 (ECIS).

[25] Shen, M., Duan, J., Zhu, L., Zhang, J., Du, X., Guizani, M., 2020. Blockchain-Based Incentives for Secure and Collaborative Data Sharing in Multiple Clouds. IEEE J. Select. Areas Commun. 38 (6), 1229–1241.

[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] 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.

[28] Elsevier, 2020. Content - How Scopus Works. https://www.elsevier.com/solutions/scopus/how-scopus-works/content.

[29] Arnaut, C., Pont, M., Scaria, E., Berghmans, A., Leconte, S., 2018. Study on data sharing between companies in Europe, Luxembourg.

[30] Lacity, M.C., Khan, S., Yan, A., Willcocks, L.P., 2010. A review of the IT outsourcing empirical literature and future research directions. Journal of Information Technology 25 (4), 395–433.

[31] Glaser, B.G., Strauss, A.L., 1967. The discovery of grounded theory: Strategies for qualitative research. Aldine, New York, USA.

[32] Oliveira, L., Zavolokina, L., Bauer, I., Schwabe, G., 2018. To Token or not to Token: Tools for Understanding Blockchain Tokens, in: Proceedings of the Thirty Ninth International Conference on Information Systems (ICIS2018), San Francisco, USA.

[33] Al-Zahrani, F.A., 2020. Subscription-Based Data-Sharing Model Using Blockchain and Data as a Service. IEEE Access 8, 115966–115981.

[34] Fernandez, R.C., Subramaniam, P., Franklin, M.J., 2020. Data market platforms. Proc. VLDB Endow. 13 (12), 1933–1947.

[35] Freitas, A., Curry, E., 2016. Big Data Curation, in: Cavanillas, J.M., Curry, E., Wahlster, W. (Eds.), New Horizons for a Data-Driven Economy. Springer International Publishing, Cham, pp. 87–118.

[36] Rantanen, M.M., Koskinen, J., 2020. Respecting the Individuals of Data Economy Ecosystems, in: Cacace, M., Halonen, R., Li, H., Orrensalo, T.P., Li, C., Widén, G., Suomi, R. (Eds.), Well-Being in the Information Society. Fruits of Respect, vol. 1270. Springer International Publishing, Cham, pp. 185–196.

[37] Lis, D., Otto, B., 2021. Towards a Taxonomy of Ecosystem Data Governance, in: Proceedings of the 54th Hawaii International Conference on System Sciences, Hawaii, USA, pp. 6067–6076.

[38] Delacroix, S., Lawrence, N.D., 2019. Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance. International Data Privacy Law.

[39] Singh, P.K., Singh, R., Nandi, S.K., Nandi, S., 2020. Designing a Blockchain Based Framework for IoT Data Trade, in: Rautaray, S.S., Eichler, G., Erfurth, C., Fahrnberger, G. (Eds.), Innovations for Community Services, vol. 1139. Springer International Publishing, Cham, pp. 295–308.

[40] Gawer, A., Cusumano, M.A., 2014. Industry Platforms and Ecosystem Innovation. J Prod Innov Manag 31 (3), 417–433.

[41] Adner, R., 2017. Ecosystem as Structure. Journal of Management 43 (1), 39–58.

[42] Ji, G., Yu, M., Tan, K.H., 2020. Cooperative Innovation Behavior Based on Big Data. Mathematical Problems in Engineering, 1–14.

[43] McBride, K., Olesk, M., Kütt, A., Shysh, D., 2020. Systemic change, open data ecosystem performance improvements, and empirical insights from Estonia: A country-level action research study. Information Polity 25 (3), 377–402.

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.

Jan Ruben Both, B.Eng. (*1990) studies industrial engineering at TU Dortmund University and has been working as a research assistant at the Fraunhofer Institute for Software and Systems Engineering ISST in Dortmund since 2018.

Prof. Dr.-Ing. Boris Otto (*1971) is holding the Chair of Industrial Information Management at TU Dortmund University. In addition to that, he is Executive Director at the Fraunhofer Institute for Software and Systems Engineering ISST. His focus areas of research are enterprise data management, industrial data ecosystems, the digital enterprise with a special focus on logistics networks as well as business networks and business engineering.

2nd Conference on Production Systems and Logistics

A Simple And Modular Approach To Path Planning For Tractor-Trailer Robots Based On Modification Of Pre-Existing Trajectories

Tobias Recker[1], Mohammad Ehsan Matour[1], Annika Raatz[1]
[1]*Institute of Assembly Technology, Leibniz University Hannover, 30823 Garbsen, Germany*

Abstract

Tractor-trailer Mobile Robot systems consist of a nonholonomic mobile robot regarded as a tractor and several passive trailers linked to the tractor via a hinge. As these systems are in many cases more economical compared to multi mobile robot systems, they are nowadays used for the transport of various objects in the field of logistics. Hence an exact and efficient path planning algorithm is required. Due to the highly non-linear characteristics of such a system, path planning is always a challenging problem. Many path planning algorithms have already been proposed for mobile robots. While some of these approaches can solve the path planning problem for robots with multiple trailers, the solutions are usually complex, limited to a specific hardware configuration, or very computationally expensive. In this work, we present an algorithm that, although still computationally expensive in its current form, provides a very simple, extensible, and flexible solution for path planning of tractor-trailer robots. Based on a conventional global planner for mobile robots (like A* or Dijkstra), our algorithm adjusts the global path according to the dimension of the carried object so that the whole system traverses a collision-free path. We take advantage of the fact that the global planner has already planned a collision-free path for the mobile robot (tractor), which the trailer follows almost exactly on straight paths. Accordingly, collisions occur predominantly in or shortly after curves. We designed our algorithm to detect these particular curves and adjust the curve radius so that no collision occurs. This way, we do not need to re-plan the whole trajectory. Also, we can upgrade almost any planning algorithm for mobile robots to work with tractor-trailer systems. We validate our method based on a series of simulations. First real-world experiments are also very promising.

Keywords

Tractor-Trailer mobile robots; Path Planning; Object Transport

1. Introduction

Large and heavy robots are often needed to transport large or heavy components. These robots not only cost more than smaller robots, but they also consume more power and occupy additional hall space. Besides, the oversizing of robots leads to a waste of raw materials and energy. In this context, it becomes necessary to consider how to increase the capabilities of smaller robots. One way to achieve this is the cooperation of a formation of smaller robots. Robot formations are often used to extend the capabilities of a single robot or to break down complex tasks into simpler subtasks [1]. In this way, formations consisting of multiple mobile platforms equipped with manipulators can transport or assemble objects that would otherwise be too heavy, big or delicate for a singular mobile manipulator. In these formations, the mobile manipulators generally consist of a nonholonomic mobile platform equipped with an industrial robot. While multi-robot formations increase a robotic system's total carrying capacity, they still require additional expensive hardware to scale to a new task. The control of these formations also increases the complexity of the transport process.

Consequently, in this work, we present an approach to replacing a robot formation member with a passive auxiliary device. The passive auxiliary device supports the transport object on one or more sides, thus reducing the robot's load. As shown in Figure 1, only a part of the object has to be held by the robot, which halves the effective gravity force.

Figure 1: Top view of a mobile manipulator (MuR 205) transporting an aluminium construction profile in combination with a roller board

In this work, we use a conventional roller board (RB) as an example of a cheap passive auxiliary device. The robot carries the RB on its loading platform and then places it at one of the object's ends before loading the transport object (see Figure 2). The object is detected using Canny Edge Detection from the Open Source Computer Vision (OpenCV) library. In our tests, the transport object was a 3 m long aluminium construction profile. One end of this profile is laid on the RB (see Figure 1). The profile can be loaded by the robot itself (see our previous work in [2]) or by a second mobile robot (when a second robot has to be used for loading, it can still perform another task during the object transport). After the robot has loaded the first end and grasped the other, it pulls the passive RB behind it as a trailer. The resulting system resembles a tractor-trailer robot (TTR), where the mobile manipulator acts as the tractor.

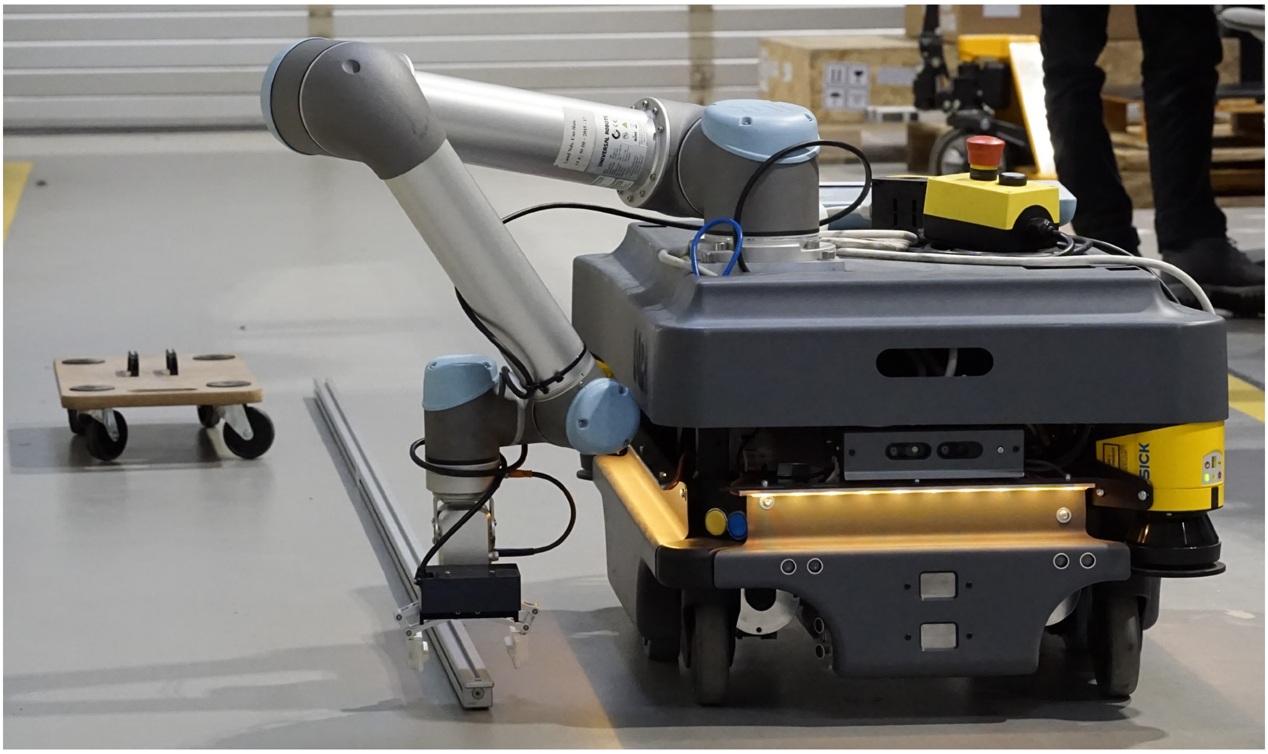


Figure 2: Top view of the TTR (MuR 205) transporting an aluminium construction profile

The main advantage of a TTR is that pulling the object on a trailer requires less energy than carrying the same object. Also, the resulting force load is much lower. Therefore, we can achieve nearly the same payload as a multi-robot system with a lower number of robots. A major disadvantage of this approach is that the RB and, thus, the object's position can no longer be actively controlled. Instead, the RB's trajectory is a consequence of the robot's trajectory, which means the RB is likely to collide with an obstacle. Conversely, the tractor's path planning must be modified to take the trailer into account.

After giving a short overview of related work, we present our modified path planning algorithm and our simulation model in section 3. In section 4, we present our simulation results and provide a conclusion and outlook on future work in section 5.

2. Related work

The problem of path planning for TTRs has been in the focus of many researchers over the last decades [3]. The various researchers have created many planners to solve path planning in different ways with different advantages and disadvantages. Although the research focuses on various aspects of the aforementioned problem, the most basic task is always to find a trajectory on which the tractor and trailer arrive at the target position in the correct orientation. This task is then made more difficult by including obstacles [4],[5], driving in reverse [6],[7] or adding more trailers [8],[9]. The first step in planning a trajectory is usually model design. A kinematic or dynamic model of the entire system is established, and subsequently, the parameters of this model are identified [10]. Researchers may choose a kinematic model over a dynamic model because it has fewer parameters and is easier to identify and compute [10]. But even for a kinematic model, the nonholonomic constraints and relatively large dimensional state-space can make modelling and identification very challenging [11].

Consequently, many approaches have been developed to deal with model design and model identification [3] [11]. These works have in common that the model is used to generate a motion plan after the modelling phase. The motion plan is then given to a controller, which tries to keep the trailer on this target by controlling the tractor [12]. This classic approach has certain advantages. For example, the position of several trailers can be controlled very precisely [9] so that the trailer(s) can be placed in the correct position for loading and unloading. This is important when, as in the case of a truck, the trailer does the actual transport, and the towing vehicle cannot unload the cargo. As we have proved in another publication [2], our system, on the other hand, is able to monitor the position of the trailer and to load and unload the transported objects itself. This means that in our case, it is not necessary to control the exact path and final position of the trailer as long as there is no collision with obstacles.

We intend to use this advantage to develop a novel approach to the path planning of tractor-trailer systems. As described, the goal of this approach is not to maximize the tracking performance or to enable complicated manoeuvring or turning procedures. Instead, we are creating an approach that is as simple as possible and can be added modularly to an existing path planner.

Other distinctive characteristics

Typical areas of application for TTRs are logistics [13],[14] and agriculture [15]. In such cases, there is a fixed coupling point for further trailers at each tractor and trailer. This coupling point is usually designed to define the coupling position precisely while the orientation is unrestricted. This allows effortless and inexpensive attachment of trailers as well as their exact control. However, this design is also very limiting. The manoeuvring behaviour of the TTR depends very strongly on the position of the coupling point [16]. Furthermore, the rigid coupling leads to a reciprocal influence of both system's error dynamics. Disturbances acting on the tractor are transferred to the tailer and vice versa. Using a manipulator to couple the object to the tractor, we can shift the coupling point's position as desired. This means we can influence the kinematic behaviour to adapt the TTR to changes in the environment. Also, the manipulator can be operated in impedance control mode, which partially dampens external disturbances.

Another difference from most existing work is that there is no persistent link between tractor and trailer. The object is only coupled to the tractor by gripping it using a robot. The connection between the RB and the object is only caused by friction. This makes the entire system very flexible in terms of the transport objects' size and shape, but the additional degrees of freedom also make it very difficult to describe the system behaviour analytically. Therefore, in the following, we present an approach that largely avoids analytical modelling and identification for path planning and instead intuitively adapts the tractor's path similar to a human truck driver.

3. Path planning algorithm

As described, the main task in literature is usually to plan the optimal trajectory for a given set of constraints. Setting up these constraints and then optimizing the path can make this process very cumbersome. Otherwise accessible objectives, such as energy or time optimized path planning or the coordination of several TTR systems to avoid collisions between robots, become very complex. This increase in complexity is caused by the coupling of kinematic constraints in a TTR. Therefore, our approach is based on a multi-step procedure (see Figure 3), which reduces the effects of coupling and enables the use of any global path planner made for a single mobile robot.

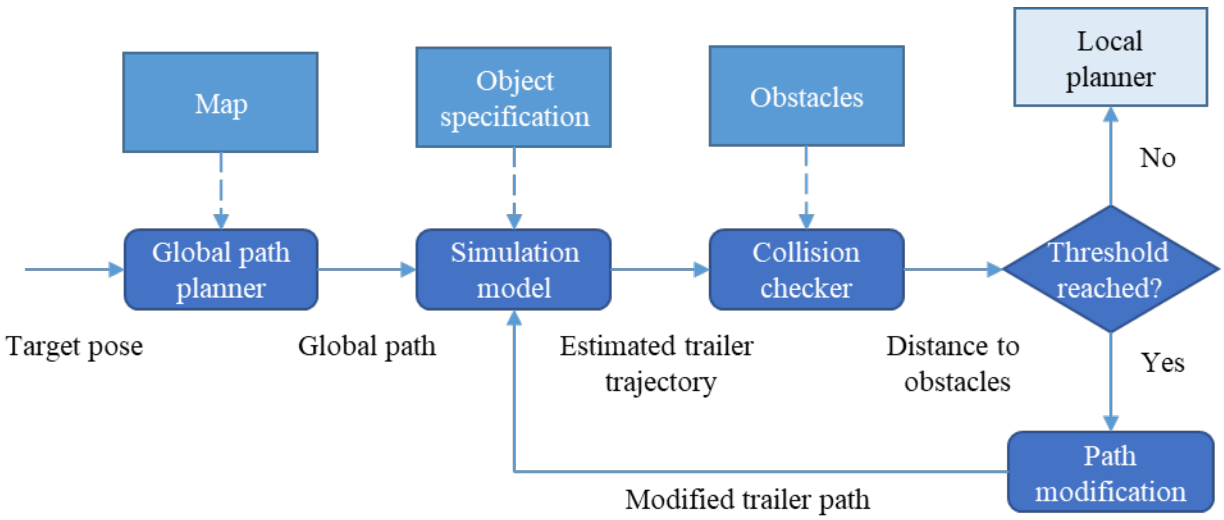


Figure 3: Flow chart of our proposed algorithm for path modification

First, we use the global path planning algorithm to generate the path target for the tractor. As we only consider the tractor in this first planning step, we can use any arbitrary planner from literature. This also means that our modification algorithm can be modularly "attached" to any existing navigation stack . To obtain an approximation for the trailer's path, we use a robot model in the Gazebo simulation environment (see section 3.1). If the simulation shows that the trailer maintains a safe distance from obstacles, we do not make any changes to the path and pass it directly to the local planner. Alternatively, we perform the procedure shown in Figure 3 to adjust the tractor's trajectory online, which we will describe in the following.

3.1 Simulation model

The simulation model is used to estimate how the trailer moves given the tractor's path. It consists of a mobile platform model, a manipulator model, an object model, and a trailer model. The models were created in the simulation environment Gazebo and are shown in Figure 4. Gazebo uses a numerical dynamics simulation to estimate the RB's motion. The advantage to an analytical model is the ease of model changes. Compared to a conventional TTR, we have to account for various additional parameters like manipulator pose, the angle of the profile, and 4 loose wheels (the typical configuration is 2 loose and 2 fixed wheels). With a numerical dynamics simulation, we can add these parameters to the model without significantly increasing complexity.

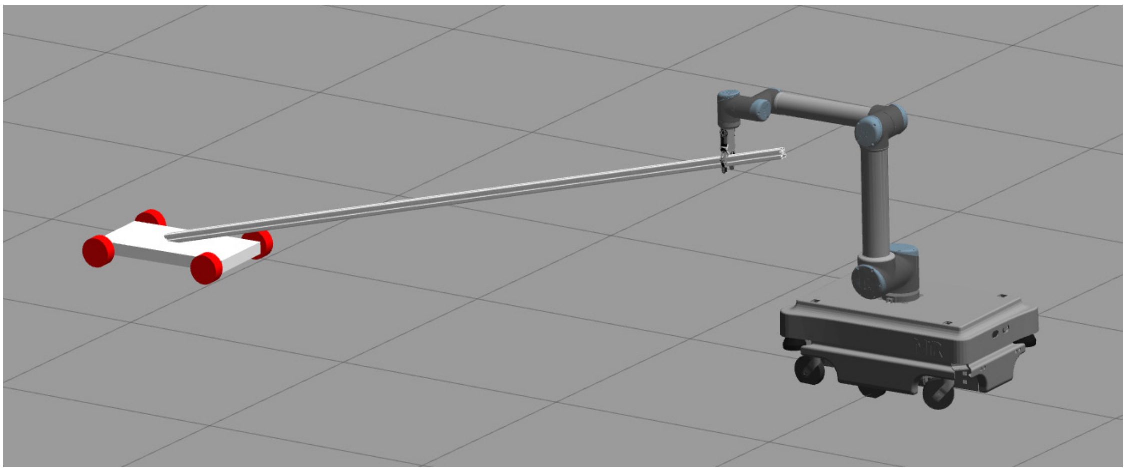


Figure 4: Model of the TTR in the simulation environment Gazebo

To build our entire system model, we used the manufacturer's existing models for the manipulator and the mobile platform. In this way, we achieve behaviour that is as realistic as possible. We calculate the object's mass, inertia, and center of gravity based on the manufacturer's CAD data and specifications. We assume that the object and its coupling to the manipulator are rigid. The coupling to the RB is made by static friction. We assume a friction coefficient μ_H of 0.5 for aluminium on wood according to [17]. The RB is modelled as a rigid plate with a length of 59 cm and a width of 29 cm. The four wheels are mounted on bearings along their vertical and rotational axes[1]. The bearings are assumed to have a friction coefficient of $c_{R,b} = 0.01$. The wheels' dimensions and distance are taken from the real roller board used to verify our simulations. The rolling resistance of the plastic wheels $c_{R,w}$ is assumed to be 0.01.

3.2 Path modification

While moving in a straight line, the trailer is always located exactly behind the tractor, eliminating the possibility of collisions with the surrounding area. Conversely, this means that all collisions with the environment occur in or shortly after curves. To minimize the computational effort, we only adjust the global trajectory segments in which the tractor takes a curve. The identification of these segments is described in section 3.2.1. Section 3.2.2 describes the modification process we have developed to adapt the existing path to a tractor-trailer system.

3.2.1 Curve detection

A curve is characterized by the presence of both linear v_x and angular velocities ω at the same time. Therefore, the first step is to derive the global path and exclude all segments, where one of the two velocities is zero. All segments where $|v_x| \gg |\omega_z|$ are also excluded, as curves with a very low curvature do not lead to any significant displacement of the trailer. Next, the start- and endpoints of the individual curve segments are determined. The start point is always the first point of the segment, and the endpoint the last point. Figure 5 shows the segmented path. The start- and endpoints of the respective curve are marked in step 3 and 4. By comparing the start orientation ϑ_S and end orientation ϑ_E, the direction of the curve can be determined in the last step. If $\vartheta_S > \vartheta_E$ it is a left curve, otherwise a right curve.

[1] For visual reasons, the wheels in our model are attached on the side instead of underneath. The kinematic behaviour is still equal.

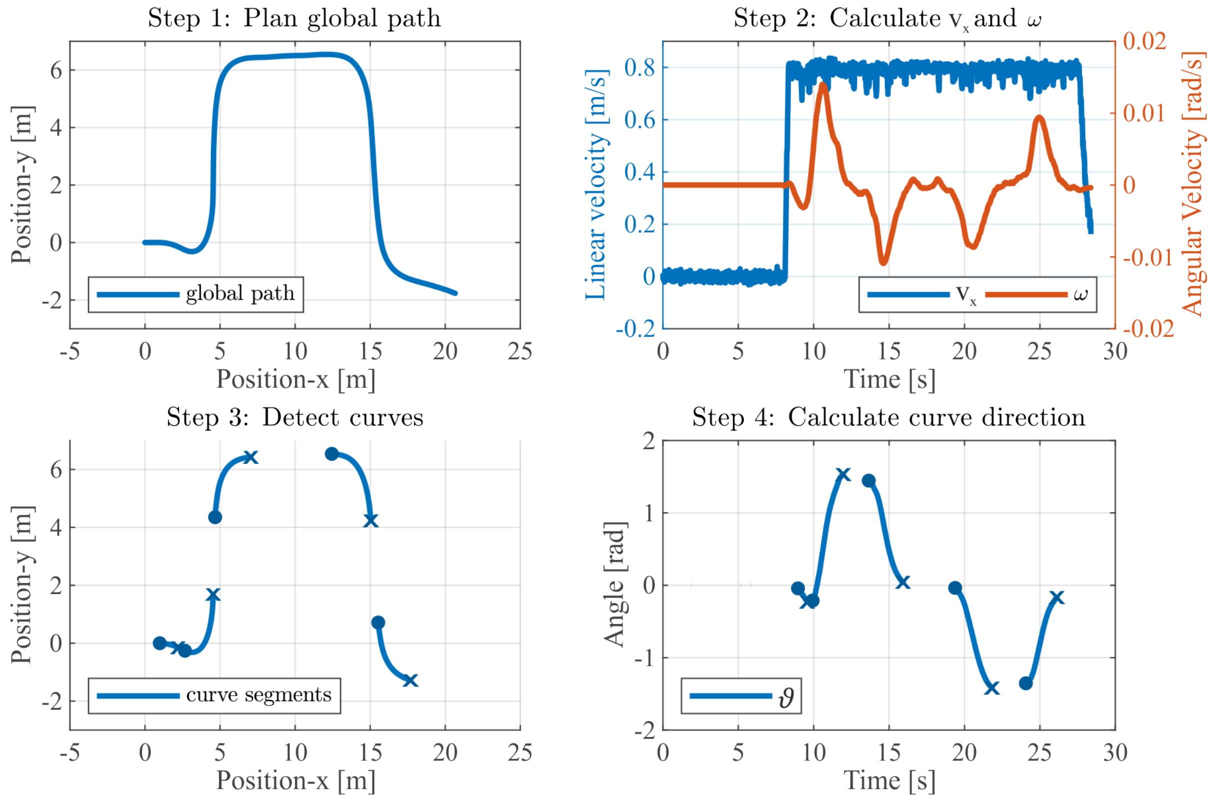


Figure 5: The four steps of curve identification. Start points are marked with a circle; endpoints with an x.

After determining all curves, which may cause a collision, the tractor's path is now modified in these segments.

3.2.2 Curve modification

In the curve modification, the simulation is used to determine the RB's minimum distance from any obstacle in every curve. If the minimum distance is below a predefined threshold, this curve is modified. We have set the threshold to 30 cm since this value is also preset as the mobile robot's safety distance. Should the RB get closer to an obstacle at the inner side of a curve, we simply increase the curve radius. This is done by defining two points $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$. Given the start pose $\boldsymbol{x}_S$ and the end pose $\boldsymbol{x}_E$ as

$$\boldsymbol{x}_S = (x_S, y_S, \vartheta_S)^T \qquad \boldsymbol{x}_E = (x_E, y_E, \vartheta_E)^T \tag{1}$$

$\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are defined as:

$$\boldsymbol{p}_1 = \begin{pmatrix} \cos(\vartheta_S) \cdot d_1 - k \cdot \sin(\vartheta_S) \cdot d_2 + x_S \\ \sin(\vartheta_S) \cdot d_1 + k \cdot \cos(\vartheta_S) \cdot d_2 + y_S \end{pmatrix} \qquad \boldsymbol{p}_2 = \begin{pmatrix} -\cos(\vartheta_E) \cdot d_1 + k \cdot \sin(\vartheta_E) \cdot d_2 + x_E \\ -\sin(\vartheta_E) \cdot d_1 - k \cdot \cos(\vartheta_E) \cdot d_2 + y_E \end{pmatrix} \tag{2}$$

The length d_1 defines the distance from $\boldsymbol{p}_1$ tangential to $\boldsymbol{x}_S$ and $\boldsymbol{p}_2$ tangential $\boldsymbol{x}_E$. The length d_2 denotes the parallel distance to the tangent through $\boldsymbol{x}_S$ and $\boldsymbol{x}_E$ where k equals:

$$k = \begin{cases} -1 \ for \ \vartheta_S < \vartheta_E \\ 1 \ for \ \vartheta_S > \vartheta_E \end{cases} \tag{3}$$

In the next step, we use a 3rd order Bézier curve to compute a new trajectory between $\boldsymbol{x}_S$ and $\boldsymbol{x}_E$. The Bézier curve is described by:

$$x(t) = \sum_{i=0}^{n} \binom{n}{i} t^i \cdot (1-t)^{n-i} \cdot b_i \tag{4}$$

with $= 3$, $t_S \leq t \leq t_E$, $\boldsymbol{b}_1 = \boldsymbol{x}_S$, $\boldsymbol{b}_2 = \boldsymbol{p}_1$, $\boldsymbol{b}_3 = \boldsymbol{p}_2$ and $\mathbf{b}_4 = \boldsymbol{x}_E$. As shown in Figure 6, the curve radius is increased depending on the parameters d_1 and d_2.

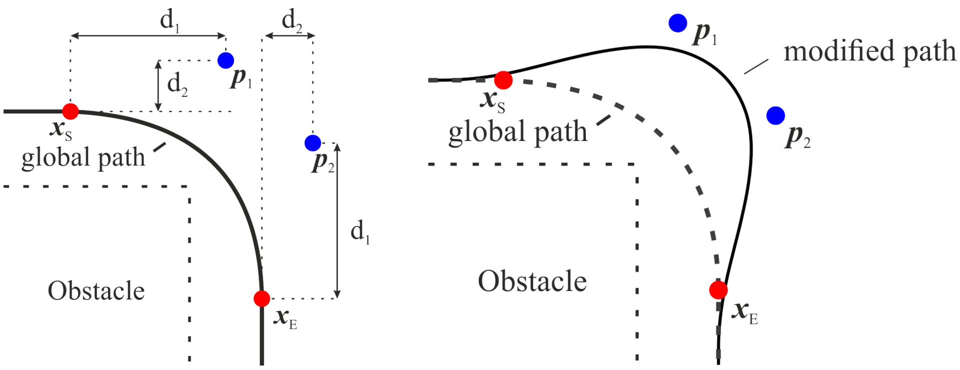


Figure 6: Bezier Curve with all control points (left: original; right: modified)

Several factors influence the selection of these parameters. On the one hand, a high value for d_1 or d_2 may result in collisions with obstacles at the outer edge of the curve. Also, the total path becomes longer, resulting in longer transport times. On the other hand, a value that is too small does not solve the original issue. Experiments show that we can achieve a good middle ground for $d_1 = l/2$, where l is the total length of the TTR. The length d_2 is determined using the local costmap. The local costmap is provided by the Robot Operating System (ROS) and represents the cost of traversing different areas of the map close to the robot. The cost map specifies, among other things, the smallest distance that the robot must maintain from an obstacle to avoid colliding with it (see Figure 8). By placing $\boldsymbol{p}_2$ at the boundary of the costmap (the next increase in cost), we can maximize the curve radius without risking a collision of the curve's outskirt. For this purpose, the smallest distance to the local costmap in the direction of $\boldsymbol{p}_2$ is determined for $d_2 = 0$. This distance is then set as d_2, whereby the maximum value is limited to $d_2 = d_1$. This process is carried out for all curves with a risk of collision. If the endpoint of a curve is closer from the start point of the next curve than d_1, both curves are combined. For this purpose, the end pose of the first curve $x_{E,1}$ and the start point of the second curve $x_{S,2}$ are deleted. The Bézier curve is then fitted with $\boldsymbol{b}_1 = \boldsymbol{x}_{S,1}$, $\boldsymbol{b}_2 = \boldsymbol{p}_{1,1}$, $\boldsymbol{b}_3 = \boldsymbol{p}_{2,1}$, $\boldsymbol{b}_4 = \boldsymbol{p}_{1,2}$, $\boldsymbol{b}_5 = \boldsymbol{p}_{2,2}$ and $\mathbf{b}_6 = \boldsymbol{x}_{E,2}$.

Lastly, the local planner uses the modified path to calculate the motion commands for the tractor. An exemplary modified global path is shown in Figure 5 on the right. This publication focuses on the global planner, so we will not describe the modifications to the local planner in detail. At this point, it is only important that the local planner uses the global trajectory, the costmap and the footprint of the robot to generate the motion commands for the tractor. We provide more detail on the estimation of the robot footprint in [16]. With this, our algorithm is complete, and we will evaluate it in the next section.

4. Evaluation

To validate our approach and as a proof of concept, we conducted simulative and physical experiments. In both cases, it was shown that collisions could be avoided even in confined environments. Nevertheless, we will focus exclusively on simulation results in this publication since they provide a much simpler and more accurate evaluation of the individual trajectories and different environments that can be tested. For the simulation, we use the model presented in section 3.1. The object is a 3-meter long aluminium construction profile with a weight of 4.5 kg. Considering the length of the profile and the gripper's additional weight, one mobile manipulator could not transport this profile alone. Three of the maps we tested are shown in Figure 7. The original global path (target path) is also marked, as created by the lattice motion planner [18].

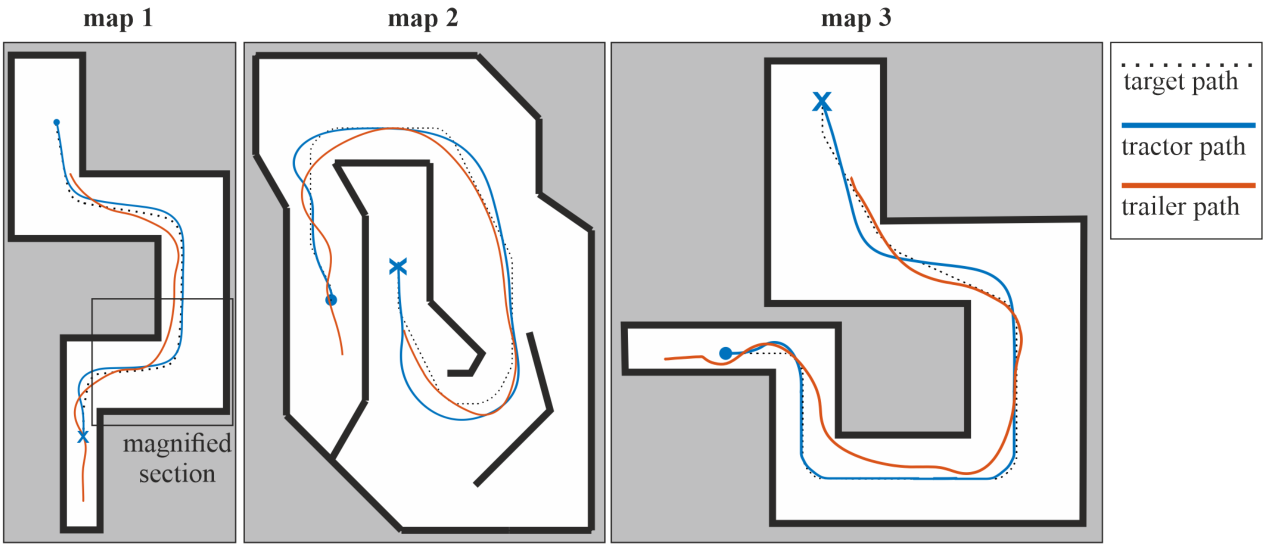


Figure 7: The three maps used for evaluation. The start of the global trajectory is marked with a circle and the end with an x.

The global path was set as the target path for the TTR and then modified as described in section 3.2. Looking at the tractor path compared to the target, both are very similar on straights and light curves. This is expected, as we only modify curves with high curvature. In modified curves, the modification results in a larger curve radius. The larger radius is especially noticeable at the end of the tractor path in map 2. Here, two curves with a high curvature are positioned closely, which causes our planner to combine both curves into one (see section 3.2.2). The increase in the curve radius also leads to an extension of the path length, as described previously. In map 1, the path length increases from 35 m to 36.2 m, in map 2 from 42 m to 45.1 m, and in map 3 from 40 m to 42.2 m. Figure 7 shows that the path modification works as intended. No collisions occur between the RB and the environment, although it appears as if the minimum distance threshold has been violated in some curves. However, this is only a visualization issue. As shown in the magnified section from map 1 (Figure 8), the higher cost region is not entered. In summary, this means that our approach is capable of generating a collision-free trajectory for a TTR.

The main disadvantage of our approach arises from the calculation of the trailer trajectory via numerical simulation. Since the calculations are currently performed on the low-powered robot computer, they can only be performed with a real-time factor (RTF) of one. This means that the RB trajectory calculation with simultaneous collision checking takes roughly the same amount of time as traversing the trajectory. However, using faster external computation, the simulation could be performed at much higher RTFs in the future.

Regarding the simulation accuracy, we observed that qualitatively the calculated trailer path corresponds to the physical trajectory. As mentioned previously, we use the manipulator to couple the trailer to the tractor. The manipulator uses an impedance controller to allow for force-free rotation of the profile around the coupling point. However, unlike the simulated manipulator, the real manipulator cannot control the rotational stiffness and rotational damping exactly to zero. This means that a small amount of stiffness and damping remains, which counteracts the RB's rotation around the coupling point. Therefore, the real profile does not rotate as much as in the simulation, causing the RB to be slightly steadier and have a slightly larger curve radius. After the first tests, it looks like this difference could be mitigated by altering the simulation model and not requiring a change in path modification.

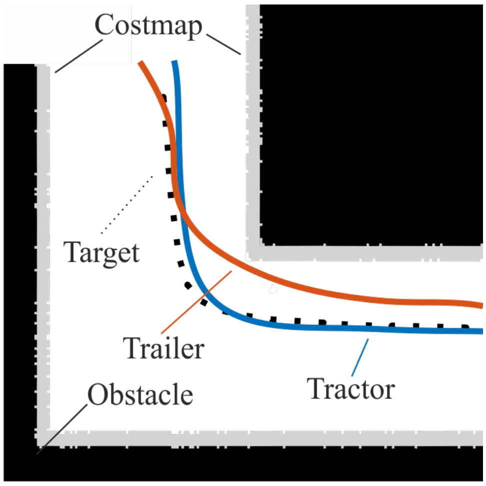


Figure 8: Magnified section from Figure 7 (map 1)

5. Conclusion and outlook

We have presented an approach that extends a single mobile robot's capabilities regarding object size and weight using an ordinary roller board. The addition of a passive RB introduces several issues for navigation and manoeuvring confined environments. To overcome these issues without significantly increasing complexity, we have presented an intuitive approach to modify an existing path that takes inspiration from a human truck driver. For this purpose, our algorithm is inserted as a module between the existing global and local planner. The modification works by first detecting all curves of the path. A simulation model then determines in which curves a collision potentially occurs. Next, we use a Bézier spline to increase the respective curve radius. This approach allows the separation of global path planning and kinematic/dynamic modelling, thereby simplifying the planning process and model design. We tested this approach in various scenarios and demonstrated its effectiveness in simulation. The transport of a construction profile shows that even minor modifications to the path can significantly reduce collision risk. For future work we plan to develop a visual system for tracking the trailer. In this way, we can perform a quantitative evaluation of physical experiments. We are also looking for a suitable analytical model. Currently, the time-consuming trailer trajectory estimation is the biggest disadvantage of our approach. The estimation time could be drastically lowered if an analytical model is established for the specific system.

References

[1] Feng, Z., Hu, G., Sun, Y. et al., 2020. An overview of collaborative robotic manipulation in multi-robot systems. Annual Reviews in Control (49), 113-127.

[2] Recker, T. Heilemann, F., Raatz, A. 2021. Handling of large and heavy objects using a single mobile manipulator in combination with a roller board. Procedia CIRP (97), 21-26.

[3] David, J., Manivannan, P.V., 2014. Control of truck-trailer mobile robots: a survey. Intel Serv Robotics (7), 245–258.

[4] Bai, G., Liang, C., Meng, Y. et al., 2019. Obstacle Avoidance of Semi-Trailers Based on Nonlinear Model Predictive Control.World Electr. Veh. J. (10), 72

[5] Yue, M., Wu, X., Guo, L. et al., 2019. Quintic Polynomial-based Obstacle Avoidance Trajectory Planning and Tracking Control Framework for Tractor-trailer System. Int. J. Control Autom. Syst. (17), 2634-2646.

[6] Pradalier, C., Usher, K., 2007. A simple and ecient control scheme to reverse a tractor-trailer system on a trajectory. International Conference on Robotics and Automation, 2208 - 2214.

[7] Jing, J., Maroli, J. M., Salamah, Y. B. et al., 2019. Control Method Designs and Comparisons for Tractor-Trailer Vehicle Backward Path Tracking. American Control Conference, 5531-5537.

[8] Pealoza-Mendoza, G. R., Hernndez-Mendoza D. E., Aranda-Bricaire, E., 2011. Time-varying formation control for multi-agent systems applied to ntrailer configuration. International Conference on Electrical Engineering, Computing Science and Automatic Control, 1-6.

[9] Michałek, M. M., 2020. Trailer-Maneuverability in N-Trailer Structures. IEEE Robotics and Automation Letters (5), 5105-5112.

[10] Wang X., Liu J., Peng H., 2021. Path Planning for Tractor-Trailer System. In: Symplectic Pseudospectral Methods for Optimal Control (97), 165-178.

[11] Ljungqvist, O., Evestedt, N., Axehill, D. et al., 2019. A path planning and pathfollowing control framework for a general 2trailer with a carlike tractor. In: Journal of Field Robotics, 1345 -1377.

[12] Binh, N.T., Tung, N.A., Nam, D.P. et al., 2019. An Adaptive Backstepping Trajectory Tracking Control of a Tractor Trailer Wheeled Mobile Robot. Int. J. Control Autom. Syst. (17), 465-473.

[13] Paszkowiak, W., Bartkowiak, T., Pelic, M., 2019. Kinematic model of multiple trailers on a tractor system for production logistics applications. Archives of Mechanical Technology and Materials (39), 16-25.

[14] Balcerak, E., Weidenfeller T., Zobel, D., 2007. Assessing the maneuverability of tractor trailer systems in heavy goods transport. Intelligent Vehicles Symposium, 511-516.

[15] Santos, L. C., Santos, F. N., Solteiro Pires, E. J., 2020. Path Planning for ground robots in agriculture: a short review. International Conference on Autonomous Robot Systems and Competitions, 61-66.

[16] Yuan,. J, Huang, Y., Sun, F., 2004. Design for Physical Structure of Tractor-Trailer Mobile Robot. International Conference on Robotics and Biomimetics, 511-516.

[17] Hering, E., Stohrer, M., Martin, R., 1995. Physikalisch-technisches Taschenbuch, VDI Volume 2.

[18] McNaughton, M., Urmson, C., Dolan J. M. et al., 2011. Motion planning for autonomous driving with a conformal spatiotemporal lattice. International Conference on Robotics and Automation, 4889-4895.

Biography

Tobias Recker (*1991) has been a research assistant at the Insitute of Assembly Technology (match) at the University Hanover since 2017. Tobias Recker is reasearching multi-robot-collaboration as well as human-robot-collaboration. His main research topic is the design and implementation of multi-robot object manipulation using mobile robots.

Mohammad Ehsan Matour (*1990) is a student at the Insitute of Assembly Technology since 2016. Mohammad Ehsan Matour received the B.Sc. degree from the University of Najafabad, Isfahan, Iran, and the M.Sc. degree from Leibniz University Hannover. His research topics include modelling and control of mechanical systems and mobile robotics systems.

Annika Raatz (*1971) has been the head of the Institute of Assembly Technology (match) at the Univerity Hanover since its establishment in 2013. Prof. Dr.-Ing. Annika Raatz is a member of the Society for Assembly, Handling and Industrial Robotics (MHI) and the Academic Association for Production Technology (WGP). Her main research topics are robot-assisted handling processes, machine concepts for handling, assembly and production automation and soft material robotic systems.

2nd Conference on Production Systems and Logistics

Hierarchical And Flexible Navigation For AGVs In Autonomous Mixed Indoor And Outdoor Operation

Maximilian Zwingel[1], Lukas Baier[1,2], Andreas Blank[1] , Jörg Franke[1]

[1]*Institute for Factory Automation and Production Systems (FAPS), Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany*

[2] *Technische Hochschule Ingolstadt, Germany*

Abstract

Globalized markets require a high degree of flexibility in existing production and logistics environments. Flexible intralogistics systems are one key component for enabling versatile production by ensuring material supply and providing a dynamic link between different production stages. Challenged by historically grown production layouts and the increasing need of adaptable supply routes due to new forms of workshop organization like matrix production, state-of-the art approaches for intralogistics systems based on automated guided vehicles (AGVs) are not sufficient for mixed indoor and outdoor operation. In order to enable a truly flexible operation of systems in these mixed indoor and outdoor scenarios, new solutions for navigation and planning as well as increased autonomous capabilities of AGVs need to be focused (e.g. adaptive detection of driveable regions).

In this paper, we propose an approach of hierarchical navigation utilizing a node-based representation of the production and logistics environment to allow flexible pathing and routing of AGVs. Based on inherent information of each node, autonomous capabilities of AGVs are activated according to the requirements of the area of operation. To prove reliability of the proposed hierarchical and flexible navigation, an evaluation of the approach is performed utilizing an industrial mobile system enhanced by autonomous capabilities in varying real-world cross-building scenarios.

Keywords

Autonomous Two Truck Systems; Network Based Pathing and Routing; Digitalized Production Environment

1. Introduction

Modern production concepts place high demands on the flexibility and availability of intralogistics systems as a link between general material supply and the individual production stages [1]. As a reactive supply system with high availability, automated guided vehicles (AGVs) are gaining popularity in an increasing number of application areas, [2] driven in particular by the transfer of high-performance sensors and algorithms from the field of autonomous driving [3] as well as advances in the area of free navigation without infrastructure. [4–7]

One key element that prevents the use of AGVs in production environments is the need for the intralogistics system to be operated between buildings with mixed traffic, as required by historically grown infrastructure with physically separated warehouses and shop floor buildings as shown exemplary by the Audi plant in Ingolstadt in figure 1. Indoor and outdoor areas place specific and different demands on navigation of AGVs, which result from different environmental influences, the volatility of the working environment and operational specifications. This concerns both the global level of navigation, involving the path planning between a starting point and a destination, and path execution (local level of navigation) including the sensing of and reaction to obstacles in the environment on the planned path. For indoor areas with relatively constant environmental conditions and quasi-static settings, different approaches to path planning have to be pursued than for outdoor areas with changing environmental conditions and frequent changes in scene due to mixed traffic. The situation is similar in the area of path execution, where the reliability and information content of different sensors differ significantly for indoor and outdoor scenarios, requiring efficient switching of the chosen solution for navigation.

Figure 1: Historically grown factory layout with mixed indoor and outdoor areas given by example of the AUDI factory in Ingolstadt in Germany. Hypothetical production areas (green) are separated from storage areas (blue) by outdoor paths of varying lengths (orange) (left: Orthofoto DOP80 © Bayrische Vermessungsverwaltung, CC BY 3.0 DE, right: © OpenStreetmap contributors, CC BY-SA 4.0)

To enable a holistic navigation of AGVs without losing the advantages of specific algorithms, we present a framework for hierarchical navigation using a node-based representation of the working environment. Introducing a top-level planning instance enables flexible navigation of AGVs and offers the possibility to activate required autonomous capabilities of AGVs by inherent information coding.

In the following chapter work related to navigation in mixed indoor and outdoor areas for AGVs is outlined. In the third chapter the proposed framework for hierarchical navigation utilizing a node-based representation of the working environment with inherent information encoding is presented and the control of autonomous capabilities based on inherent information is detailed. In chapter 4 the hierarchical approach for flexible navigation is exemplary implemented using an AGV in a reference scenario for evaluation purposes. Finally, an outlook on further improvements is given.

2. Related work

Research in the field of mixed indoor and outdoor navigation for AGVs has been gaining popularity lately following the increasing availability of adequate sensors and outdoor capable vehicles as well as press coverage of companies planning to utilize autonomous robots for last mile delivery.

In the context of AGVs, the task of navigation can be divided in the three sub-tasks of

- the abstraction of the working environment,
- the planning of a global path and
- the execution of a resulting trajectory following the path.

The primary approach towards abstraction of the working environment in a mixed indoor and outdoor environment is done by extending established solutions for indoor use cases into the mixed indoor and outdoor environment. [8] describes the use of simultaneous localization and mapping (SLAM) in mixed indoor and outdoor scenarios, without addressing path planning as part of the holistic task. A hierarchical approach for the representation of the working environment using SLAM generated maps on an execution level and a connection graph on the planning level suitable for navigation purposes is described in [9]. Addressing the high demand on memory and processing power generated by a single map, a semantic approach with autonomous detection of traversing areas and triggering of corresponding SLAM configurations is shown in [10]. While the potential of SLAM for localization and map generation in static environments with a high number of references is indisputable, the applicability for holistic navigation in mixed indoor and outdoor scenarios is low. The main reasons can be found in the limited environmental information provided by maps generated using onboard sensors and the necessity for prior recording of maps to allow for autonomous navigation.

[11] addresses the problem of poor scalability using metric maps by combining a topological planning layer with multiple local SLAM maps enabling for flexible pathing while simultaneously resulting in significantly improved planning times and comparable performance to metric maps. However, this approach still depends on map generation with onboard sensors prior to autonomous navigation. Yet it shows the clear advantages of a hierarchical approach for navigation.

For pathing of AGVs in mixed indoor and outdoor environments, graph-based solutions using topological or hybrid maps show clear advantages over metric maps for large and changing environments. Combined with adequate solving algorithms, an optimal path can theoretically be found at any time with little computational demand, if all relevant factors are known to the system. [12]

Alternative approaches for unknown working environments skip global path planning and instead focus on a reactive concept for trajectory execution, by using vision sensors combined with artificial intelligence. This approach yields similar results to an uninformed path planning algorithm for changing environments. [13,14]

An approach towards a flexible navigation based on changing environmental conditions is detailed in [15] with a focus on an adaptive localization for mobile robots in indoor and outdoor environments using an Extended Kalman Filter (EKF). However, the solution is based on an already planned trajectory with inherent information to control the sensor sources accordingly and cannot be transferred to general logistic tasks.

The state of the art shows that there are multiple solutions for the abstraction of working environments as well as path planning and execution for AGVs. Using inherent information regarding the working environment in conjunction with optimized planning and execution however has not yet been investigated in the area of mixed indoor and outdoor intralogistics.

3. Hierarchical Framework for navigation of AGVs

Based on the analysis of related work it becomes apparent, that no solution yet is suitable for the flexible navigation of AGVs in mixed indoor and outdoor scenarios. The abstraction of the working environment must ensure that even large deployment environments are aggregated in a suitable way so that the corresponding storage and implementation effort before deployment is reduced. At the same time, relevant information for global path planning must be preserved. Specifically metric information as well as different environments and their influences on the operation of the vehicles have to be considered. Global path planning must be realized in a way, that all information, namely current environmental conditions, the physical setup of the AGV as well as traffic rules is taken into account. Similarly, the execution of a path has to be reactive and optimized based on the area of operation. To meet these requirements, we propose a hierarchical framework for navigation of AGVs using a node-based representation of the working environment with inherent information for global pathing and local trajectory execution, as shown in figure 2.

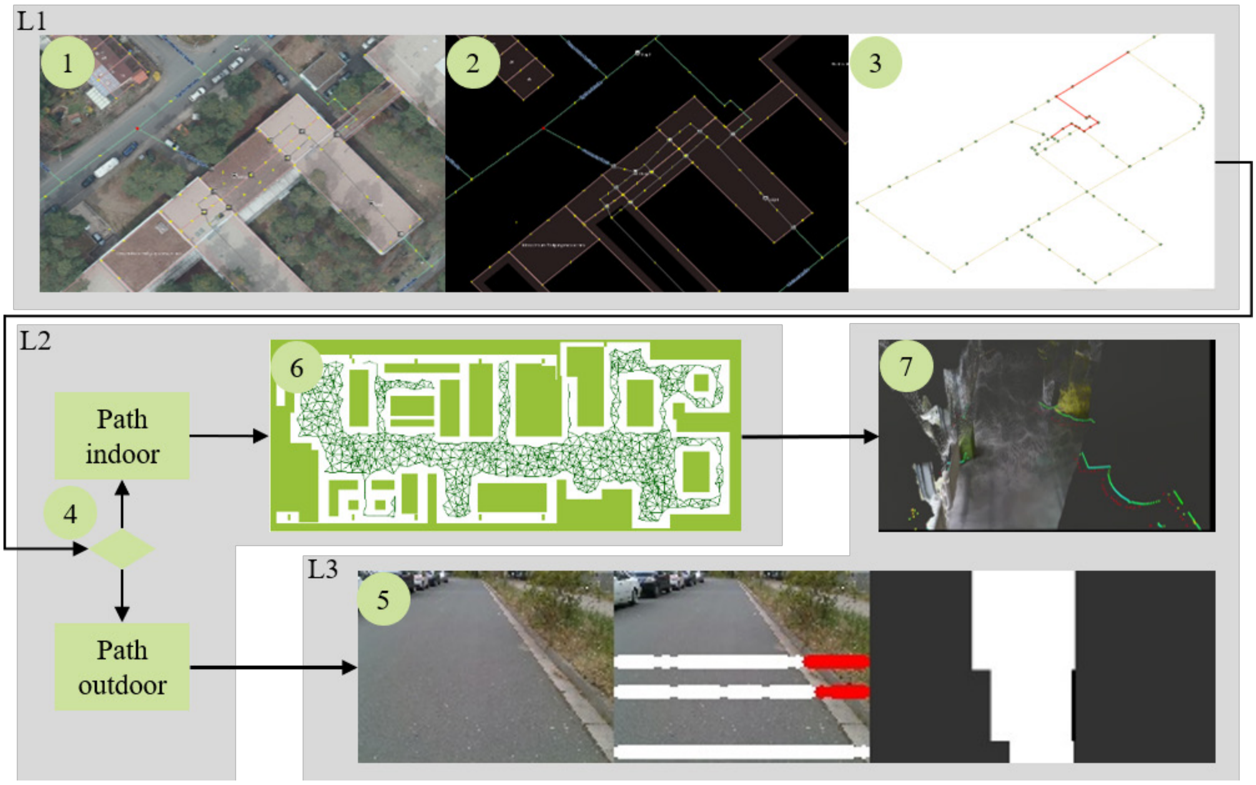


Figure 2: Hierarchical framework for navigation of AGVs in mixed indoor and outdoor operation, starting with a metric map of the whole working environment (1), a node based representation in hybrid format (2) and top level graph planning (3) followed by inherent decision making for indoor and outdoor navigation in the second layer (4) with a definition of driveable regions for outdoor areas (5) and a combination of probabilistic roadmaps (6) and local trajectory planning (7) for indoor areas.

3.1 First layer: Top level navigation using a node-based representation of the working environment

An adequate digitization of the working environment in a suitable format for flexible navigation of AGVs in mixed indoor and outdoor environments is achieved by using a node-based representation following the osm standard. Based on absolute coordinates of single points (nodes) and relations between them (ways) as well as self-defined key-value pairs for inherent information, a complete digitization of relevant information is possible. The resulting hybrid map contains a topological description of the working environment with metric information encoded in the coordinates of each node (as seen in Figure 2 Pos. 2). The map is stored in an xml-like format, which is imported into a database and evaluated predicated on defined metrics. Intrinsic information about the environment of each node (indoor, outdoor or transition) as well as restrictions in traveling speed and direction for ways are saved inside the xml file similar to a navigation approach for pedestrians presented in [16].

Utilizing the topological map on each AGV in addition to configuration files for weight manipulation based on each vehicle`s capabilities and the current environmental conditions, a directional graph with specific costs for each way is generated. This graph is then used for top level path planning using an A*-Algorithm to find the optimal top-level path between the current position of the AGV and its destination, as shown in figure 2, Pos. 3. If a reconfiguration is necessary, each node can be used as a new starting point with adjusted weights based on the changed conditions.

The list of nodes to travel for the AGV is exported as a file with comma separated values stating the number and order of nodes to travel, the coordinates of each node in GNNS format, the travel distance between each node and the environmental flag corresponding to each node. Additionally, for indoor nodes, a corresponding flag for a detailed map is given. This list is used by the second layer of navigation.

3.2 Second layer: Global path planning based on inherent information of the top-level path

The second layer of navigation is responsible for path planning in an environmentally global context, e.g. inside one production hall or connected outside areas. Based on the environmental flag of the node list inside the csv-file, the AGV is able to choose the optimal method for global path planning based on the current environmental conditions. If a set of nodes is marked being outdoor, a direct heading approach is proposed based on the circumstance, that outdoor areas often are not well structured and rapidly alternating in nature, which makes the use of complex path planning tools less reasonable. Thus, the direct heading between two nodes of the top-level graph, calculated based on the GNNS coordinates of both nodes, is sufficient and the connecting line is given as a path.

For indoor areas however, a global pathing is proposed using a probabilistic roadmap as detailed in [17]. This allows for an optimized path planning in structured indoor areas without overhead on the top level of navigation. A linkage between the nodes of the top-level navigation and the coordinates in the global navigation context is given by static transformation between the flagged map and the GNNS coordinate frame. The use of a global pathing in the second level of AGV navigation allows for an optimized routing inside buildings similar to classic indoor use cases of AGVs, while still reducing the application of said algorithms to meaningful use cases.

3.3 Third layer: Path execution in local context

The final layer of the proposed framework for navigation of AGVs in a mixed indoor and outdoor environment is path execution in the local context. As described in 3.2, for outdoor areas only a general heading towards the next node of the top-level navigation is given. Based on the configuration linked to the specific AGV and its outdoor navigation skills, corresponding autonomous capabilities are activated to achieve a reactive local navigation. For unstructured environments we propose the definition of driveable regions based on optical sensors to define the local path corresponding to the general heading, as shown in figure 3.

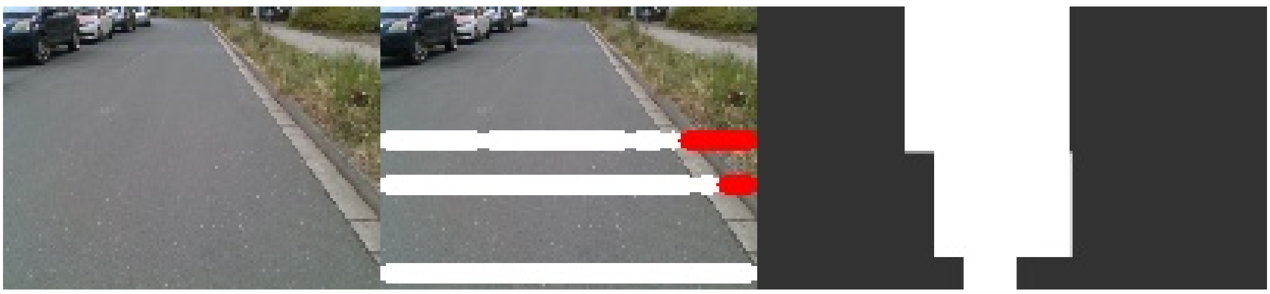


Figure 3: Definition of driveable regions based on optical sensors, showing the initial image (left) the detected driveable regions at certain ranges (middle) and the corresponding cost map for driving (right)

This definition of driveable regions ensures the operation in changing and challenging environments based on the autonomy of the system itself, similar to the capabilities of a human operator. Additionally, different configurations for image processing algorithms can be used to suit different AGVs and to consider properties like maximal height of kerbing to traverse or off-road capabilities.

For indoor areas with structured environments, standard approaches for local navigation like the Dynamic Window Approach (DWA) or Times-Elastic-Band (TEB) are suitable based on the sensors and configuration of the specific AGV. This ensures proven performance in indoor use cases with short detection ranges and a higher possibility of encounters with other road users.

4. Implementation of the proposed framework on an AGV system

To validate the proposed framework, it is implemented on an autonomous tow truck with a rugged industrial computer running Linux Ubuntu 20.04 and the Robot Operating System (ROS noetic). The node based representation of the working environment is saved as an osm-file containing the information described in 3.1. The top-level navigation is implemented utilizing a python script for loading the xml-formatted data into a PostgreSQL database with a PostGIS extension for GNNS coordinate handling and selection of the right configuration file. A second python file is executing the corresponding SQL-queries for path calculation according to the configuration file and saving of the planned path. Both python scripts are linked to an action server in ROS and can thus be executed by each AGV itself. The resulting optimal path is saved in a predefined directory and processed by an adaptive state machine addressing the evaluation and control of the second and third level of navigation. In our implementation, the software FlexBe is used for this task. Details on the functionalities and features of said adaptive state machine can be found in [18].

Global path planning and trajectory execution in the second and third layer of navigation are implemented using plugins for the standard navigation package inside ROS. According to the environment, the corresponding states inside the adaptive state machine load or unload the specific plugins for global path planning and trajectory execution inside the navigation package and monitor execution progress. Coordinate transformation between the local references of indoor maps and the GNNS coordinates is done utilizing the tf2 package and static transforms between the coordinate frames. The AGV itself is capable of autonomous navigation using different sensors detailed in [19] and is physically capable of indoor and outdoor operation.

The reference scenario for evaluation of the autonomous operation of AGVs in mixed indoor and outdoor application utilizing our proposed approach was defined at the chair of Factory Automation and Production Systems (FAPS) at the Friedrich-Alexander-University Erlangen-Nuremberg containing multiple indoor and outdoor areas connected with a manually controlled gate as shown in figure 4.

Figure 4: Reference scenario for mixed indoor and outdoor operation of an autonomous AGV using the proposed framework

5. Evaluation of the proposed framework

For evaluation purposes, transport scenarios under different environmental conditions were conducted using the implementation of the proposed approach for specified use cases. Transportation orders with mixed indoor and outdoor areas were issued and the functionality and the performance of the overall system was monitored.

As a first use case, two different configurations simulating different weather conditions for a defined transport order were issued and the behaviour of the top-level navigation was evaluated qualitatively, as seen in figure 5. The different weighting of indoor and outdoor ways for the pathing result in two different outcomes, both optimal in regards to the corresponding configuration. For weather without negative impact on safety or performance of AGVs this means outdoor traveling speeds of 8 km/h and indoor traveling speeds of 6 km/h, resulting in a travel route completely outdoors. For weather with negative impact, the outdoor speed is reduced as low as 0.001 km/h in regard to the severity of the impairment, resulting in a route mostly indoors. A speed larger than zero is needed to enable planning if no indoor route between a start and target point exists.

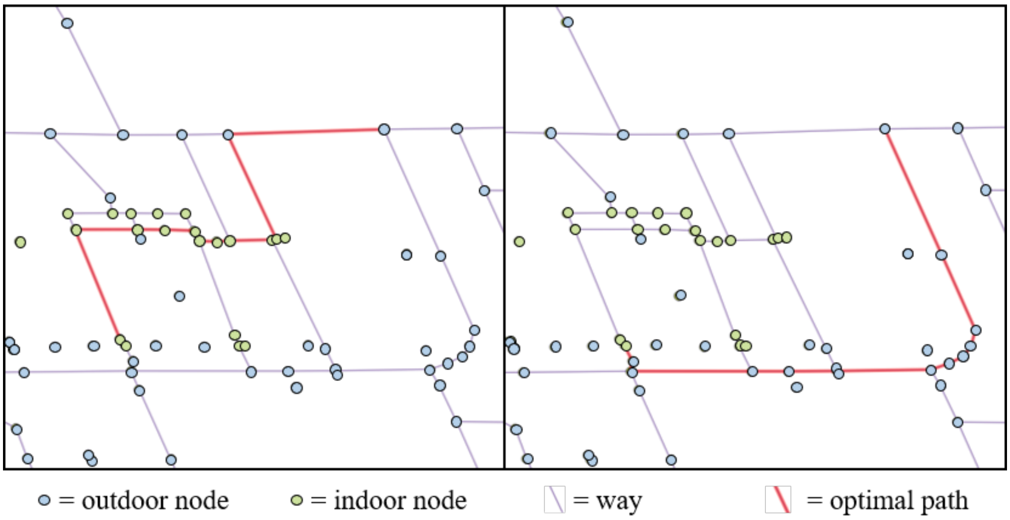


Figure 5: Results for top level navigation based on configurations simulating bad weather (left) and good weather (right)

To prove applicability to intralogistics systems, the calculation time of the approach was evaluated in addition to the functionality for the approach using normal consumer-grade hardware. Table 1 shows the needed computation time for top-level navigation in relation to the nodes inside the digitized working environment using an Intel i5-9300H CPU with 16 GB of total RAM and a SanDisk SSD. The value shows the average of 100 queries including parsing of the osm-data, path generation and saving of the corresponding data for randomly generated start and target points.

Table 1: Evaluation of calculation time depending on the number of nodes

Map file	**faps.osm**	**erlangen.osm**
Number of nodes	2113	77218
Elapsed time (s)	1.632	3.417
User CPU time (s)	0.085	0.329

Even for rather large maps, the calculation time for the top-level navigation is adequate for industrial use of AGVs in mixed indoor and outdoor environments. Performance evaluation of the different algorithms for the second and third level of navigation can be found in the corresponding publications detailed in chapter 3.2 and 3.3.

6. Summary and outlook

In this paper, we propose a hierarchical navigation for flexible use of AGVs in mixed indoor and outdoor environments. Using a node-based representation of the environment, a top-level navigation and configuration files suitable to each AGV can be used to plan optimized routes for intralogistics tasks in a holistic context. Based on corresponding flags utilizing inherent data and an adaptive state machine the use of suitable algorithms for the second level of path planning and third level of trajectory execution is enabled. The evaluation of the hierarchical approach shows the potential of the proposed framework and feasibility even for large areas of operation. However, evaluation in other mixed indoor and outdoor environments is needed for final evaluation of the limits of said framework. Additionally, the support for multiple AGV systems is possible, but not yet implemented nor tested.

Acknowledgements

The research in this paper was conducted in relation to the Project "E|SynchroBot" (20184 N, Forschungsgemeinschaft Intralogistik/Fördertechnik und Logistiksysteme e.V.) which is supported by the German ministry for Economy and Energy by decision of the German Parliament.

References

[1] Wiendahl, H.-P., ElMaraghy, 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), pp.783–809.

[2] Carsten Heerphone, 2020. SERVICE ROBOTS Record: Sales Worldwide Up 32%: Logistics Robots up 110%, Medical Robots up 28%, https://ifr.org/news/service-robots-record-sales-worldwide-up-32.

[3] Badue, C., Guidolini, R., Carneiro, R.V., Azevedo, P., Cardoso, V.B., Forechi, A., Jesus, L., Berriel, R., Paixão, T.M., Mutz, F., Paula Veronese, L. de, Oliveira-Santos, T., Souza, A.F. de, 2021. Self-driving cars: A survey. Expert Systems with Applications 165, 113816.

[4] Chang, C.-K., Siagian, C., Itti, L., 2010 - 2010. Mobile robot vision navigation & localization using Gist and Saliency, in: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2010), Taipei. 18.10.2010 - 22.10.2010. IEEE, pp. 4147–4154.

[5] Kedilioglu, O., Lieret, M., Schottenhamml, J., Würfl, T., Blank, A., Maier, A., Franke, J., 282021. RGB-D-based Human Detection and Segmentation for Mobile Robot Navigation in Industrial Environments, in: Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. 16th International Conference on Computer Vision Theory and Applications, Online Streaming. 08.02.2021 - 10.02.2021. SCITEPRESS - Science and Technology Publications, pp. 219–226.

[6] Mäkelä, H., Numers, T. von, 2001. Development of a navigation and control system for an autonomous outdoor vehicle in a steel plant. Control Engineering Practice 9 (5), pp. 573–583.

[7] Siagian, C., Chang, C.K., Itti, L., 2014. Autonomous Mobile Robot Localization and Navigation Using a Hierarchical Map Representation Primarily Guided by Vision. J. Field Robotics 31 (3), 408–440.

[8] Bosse, M., Newman, P., Leonard, J., Teller, S., 2004. Simultaneous Localization and Map Building in Large-Scale Cyclic Environments Using the Atlas Framework. The International Journal of Robotics Research 23 (12), pp. 1113–1139.

[9] Estrada, C., Neira, J., Tardos, J.D., 2005. Hierarchical SLAM: real-time accurate mapping of large environments. IEEE Trans. Robot. 21 (4), pp. 588–596.

[10] S. F. G. Ehlers, M. Stuede, K. Nuelle, T. Ortmaier, 2020. Map Management Approach for SLAM in Large-Scale Indoor and Outdoor Areas, in: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 9652–9658.

[11] Konolige, K., Marder-Eppstein, E., Marthi, B., 2011 - 2011. Navigation in hybrid metric-topological maps, in: 2011 IEEE International Conference on Robotics and Automation. 2011 IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China. 09.05.2011 - 13.05.2011. IEEE, pp. 3041–3047.

[12] Miler, M., Medak, D., Odobašić, D., 2014. The shortest path algorithm performance comparison in graph and relational database on a transportation network. PROMET 26 (1), pp. 75–82.

[13] Seiya, S., Carballo, A., Takeuchi, E., Miyajima, C., Takeda, K., 2018 - 2018. End-to-End Navigation with Branch Turning Support Using Convolutional Neural Network, in: 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO). 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), Kuala Lumpur, Malaysia. 12.12.2018 - 15.12.2018. IEEE, pp. 499–506.

[14] T. Krajník, F. Majer, L. Halodová, T. Vintr, 2018. Navigation without localisation: reliable teach and repeat based on the convergence theorem, in: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1657–1664.

[15] E. I. Al Khatib, M. A. K. Jaradat, M. F. Abdel-Hafez, 2020. Low-Cost Reduced Navigation System for Mobile Robot in Indoor/Outdoor Environments. IEEE Access 8, 25014–25026.

[16] Wang, Z., Niu, L., 2018. A Data Model for Using OpenStreetMap to Integrate Indoor and Outdoor Route Planning. Sensors 18 (7), 2100.

[17] Scholz, M., Zwingel, M., Schuderer, P., Franke, J., 2019. Sustainable Intralogistics due to Uniform Software and Modular Transport Entities. Procedia CIRP 80, pp. 239–244.

[18] Schillinger, P., Kohlbrecher, S., Stryk, O. von, 2016 - 2016. Human-robot collaborative high-level control with application to rescue robotics, in: 2016 IEEE International Conference on Robotics and Automation (ICRA). 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm. 16.05.2016 - 21.05.2016. IEEE, pp. 2796–2802.

[19] Zwingel, M., Herbert, M., Lieret, M., Schuderer, P., Franke, J., 2020. An Economic Solution for Localization of Autonomous Tow Trucks in a Mixed Indoor and Outdoor Environment Using a Node Based Approach, in: , Production at the leading edge of technology. Springer, Berlin, Heidelberg, pp. 457–466.

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.

Lukas Baier is a research associate at the Technische Hochschule Ingolstadt and head of the technology field Production control and Intralogistics. His research focusses on sensor technologies and data processing.

Andreas Blank has been a research associate at the Institute for Factory Automation and Production Systems at the Friedrich-Alexander-University Erlangen-Nuremberg and head of the technology field Industrial and Service Robotics. His research is focused on teleoperation and autonomous robots.

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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Feature Engineering for a Cross-process Quality Prediction of an End-of-line Hydraulic Leakage Test using an Experiment Sample

Christian Neunzig[1,2], Simon Fahle[1], Bernd Kuhlenkötter[1], Matthias Möller[2]

[1]Chair of Production Systems, Ruhr-University Bochum , Universitätsstraße 150 , 44801 Bochum, Germany
[2]Bosch Rexroth AG, Bexbacher Straße 72, 66424 Homburg, Germany

Abstract

The increasing availability of manufacturing data and advanced analysis tools are forcing the demand for data-driven approaches to improve the quality of workpieces and the efficiency of manufacturing processes. The analysis of real manufacturing data is challenging due to frequent changes in production circumstances. In this work, machine learning methods based on the data along the value chain of hydraulic valves are used to predict the leakage results during end-of-line testing. The leakage volume flow measurement results are very sensitive to changes in gap geometry and temperature level in the measurement cross-section. Additional measurements and experiments are required to interpret the systematic influences of the input data on the target variable and to introduce the missing information into the dataset. The design of a meta-model using experiment data supports the identification of statistical patterns to be applied to the real production dataset as a feature. This paper presents a systematic approach to hand-crafted feature engineering that improves the quality prediction of end-of-line hydraulic leakage testing.

Keywords

Feature Engineering; Hydraulics; Machine Learning; Quality Control

1. Introduction

In the last 20 years, the use of computational fluid dynamics (CFD) has produced remarkable advances and developments for optimizing geometries and understanding hydraulic phenomena [1], [2]. The application of machine learning is motivated by tremendous success in optimizing the product and process quality of industrial goods. Machine learning techniques support the continuous improvement process to reduce manufacturing costs by avoiding non-value-added production steps and to ensure product quality. [3], [4] Product quality include measures of the inner work piece state such as an internal leakage [5], [6]. The manufacturing of hydraulic control valves for agricultural peripherals consists of the machining process, assembly, and end-of-line inspection of 100% of the workpieces. In this use-case, the hydraulic testing is the most costly production step and the bottleneck in the manufacturing.

To relieve the test bench capacity, machine learning techniques are used to predict the end-of-line test results of the internal hydraulic leakage based on the cross-process data of the different series production steps. The internal leakage of the valve is intrinsic as the spool trembles at defined positions within the body to ensure the functionality of the valve, but it must be minimized due to the loss of performance [7]. The measurement results of the internal leakage are considered as labels of a supervised regression problem.

Exploiting existing manufacturing data to approximate the physical relationships of input and target variables is a promising strategy to add value from the necessary data management [3]. The requirement for minimum production cost inhibits data collection of significant thermal and geometric information. The naïve approach of measuring component and fluid temperature is neither economically nor technically feasible for two reasons. First, simultaneous temperature measurement of the relevant components in the closed body of the valve is not possible without altering the physical bodies. Second, measuring the temperature of the fluid alters the fluid dynamics within the part. The sophisticated testing set-up requires the representation of the temperature and the gap geometry in the cross-section by other quantities. Extending the features through a systematic investigation with additional measurements, similar to experimental design methods, can provide significant benefits by adding knowledge to the predictive model [8]. Due to the high effort and lack of machine availability, the amount of experimental data is small. [10] The concept of meta-learning provides a solution to overcome the limited amount of labeled data. The use of multiple source tasks for prediction causes a beneficial initialization for learning on the target task [9]. A meta-regression model is developed to investigate the relationship between the gap geometry and the representative temperature variables based on an experimental dataset.

The paper begins with an overview of related work in meta-learning and applied machine learning and continues with an introduction to the use-case. This is followed by the design of an experiment dataset that acts as a meta-training dataset. A meta-regression model derives the coefficients c_i that best fit the meta-linear regression model. The resulting feature, called the t_{factor}, is applied to the production dataset to assess its predictive power and the functionality of the procedure. The paper concludes with a conclusion section and suggestions for further work.

2. Related work

2.1 Meta-Learning

Learning from a few examples is an essential aspect of human intelligence. One possibility that enables gain solutions to complex tasks from just a few examples is to use experience to learn a prior about tasks. [24] Meta-learning is an approach to learning from imbalanced datasets that are common in the real world. The challenge is to learn an accurate meta-model when a dominant task set has only a small number of training examples. [17] Meta-learning can be divided into mixed linear regression (MLR) and multi-task learning (MTL). The number of samples, the dimension of data points, and the number of clusters are defined as n, d, and k, respectively. In MLR, the sample and time complexity are either polynomially dependent on k or depends on the inverse of the k^{th} singular value of some moment matrix. In contrast, in MTL, all tasks have the same number of examples and performance is evaluated on the observed tasks used in training. [10] There are two perspectives on approaching meta-learning: optimization based ([18], [19], [20], [21]) and probabilistic ([22], [23], [24], [25]).

According to KONG ET AL. the formulation of a standard meta-training for supervised learning, motivated by a probabilistic view captured in formulas (1)–(4), assumes a given collection of n meta-training tasks $\{\boldsymbol{\mathcal{T}}_i\}_{i=1}^{\mathrm{n}}$ derived from some distribution $\mathrm{P}(\boldsymbol{\mathcal{T}})$. [10] Each task $\boldsymbol{\mathcal{T}_i}$ is associated with a dataset of size t_i that is defined as the meta-training dataset in (1):

$$\boldsymbol{\mathcal{D}}_{meta-train} = \{\{(\boldsymbol{x}_{i,j}, y_{i,j}) \in \mathbb{R}^d \times \mathbb{R}\}_{j\in[t_i]}\}_{i\in[n]} \quad (1)$$

Scooping some structural patterns in $\mathrm{P}(\boldsymbol{\mathcal{T}})$, the goal is to train a model for a new task $\boldsymbol{\mathcal{T}^{new}}$ that is part of $\mathrm{P}(\boldsymbol{\mathcal{T}})$, from a small amount of training dataset in (2)

$$\boldsymbol{\mathcal{D}} = \{(\boldsymbol{x}_j^{new}, y_j^{new})\}_{j\in[\tau]} \quad (2)$$

Each task $\boldsymbol{\mathcal{T}_i}$ is associated with a model parameter c_i, where the meta-training data is taken from (3):

$$(\boldsymbol{x}_{i,j}, y_{i,j}) \sim P_{c_i}(y|\boldsymbol{x})P(\boldsymbol{x}) \text{ for all } j \in [t_i] \tag{3}$$

The prior distribution of the tasks and the model parameters are completely characterized by a meta-parameter θ that (4):

$$c_i \sim \mathrm{P}_\theta(c) \tag{4}$$

However, task ambiguity is a critical problem to deal with, if the goal of meta-learning is to learn solutions to new tasks from small datasets. If each task has only one sample, the problem is explored. Even with the best possible prior, there may simply not be enough information in the examples for a new task to solve that task with a high degree of confidence. Therefore, it is useful to develop meta-learning methods that can suggest multiple possible solutions to an ambiguous learning problem. Such a method can be used to evaluate uncertainty, perform active learning, or obtain direct human control over which pattern to prefer. [24] In some cases, the meta model cannot achieve vanishing error when there is noise [27].

2.2 Applied Machine Learning in Manufacturing

Machine learning methods can be distinguished by their learning strategy (e.g., supervised, unsupervised), prediction task (e.g., regression, classification), and structure (e.g., shallow, deep) [16]. The number of data points per input dimension into the model and the model complexity are strongly interrelated. For regression problems in manufacturing data with a ratio of data points to input dimensions smaller than 1000, WEICHERT ET AL. recommend the use of the following groups of machine learning algorithms with descending suitability: decision tree ensembles, SVM, linear regression, and artificial neural networks. For the selection of the suitable model, the relationship between the process complexity, the amount of data stored, and the model complexity is significant. If this is not taken into account, complex models that have been trained on small amounts of data can be trained with an increased risk of overfitting and/or lack of interpretability. [3] For this reason, the Extreme Gradient Boosting (XGBoost) is considered as a widely used and state-of-the-art representative of gradient boosting ensemble trees. [28], [29] In the sequential boosting process illustrated in Figure 1, each decision tree depends on the result of the previous tree to produce an improved prediction. The boosting algorithm is built incrementally by correcting the errors of the previous weak models. [30]

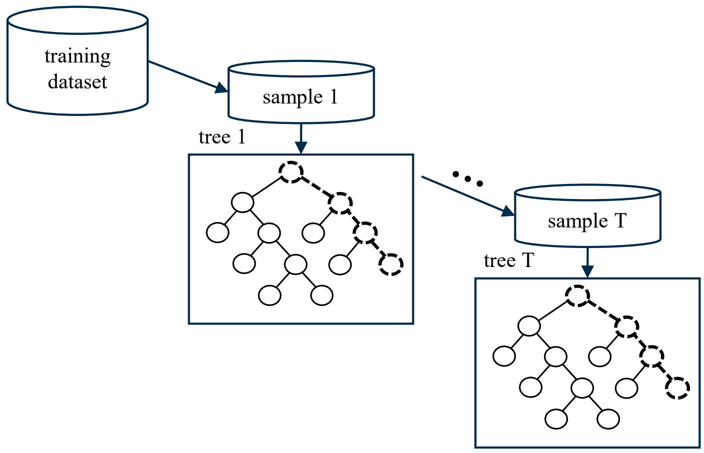


Figure 1: Boosting algorithm principle; cf. [29]

3. Use-Case: Hydraulic Leakage Flow of Directional Valves

The considered technical system is a control valve under test conditions. The electro-hydraulically actuated directional valve has four significant spool positions. The spool is actuated by a 4/3 directional pilot control valve which pressurizes the outer chambers to move the spool. The position of the control spool is fed back to the on-board electronics by means of a displacement transducer. The electronics control the position of the spool depending on the requested set point using the current from the pilot valve. [11], [12] The desired flow is set at the consumer ports to operate the connected equipment, which is simulated by a test bench. The spool is characterized by various diameters in order to shut off different pressure spaces in the four defined spool positions. In general, the cross-section under consideration is simplified as an arrangement of an eccentric spool in a housing bore with the absolute eccentricity e of the spool, as shown in Figure 2.

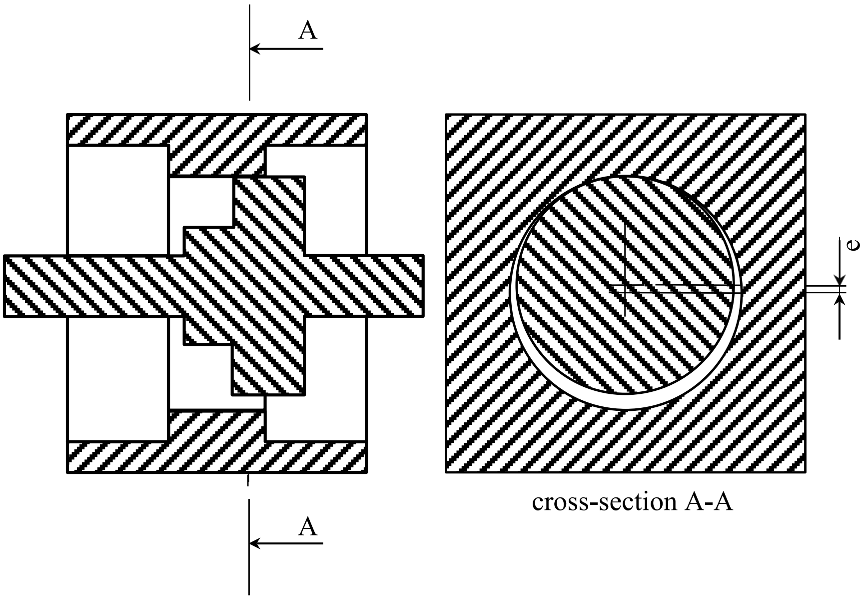


Figure 2: cross-section of an eccentric spool in a bore

Physically inevitable, the leakage flow between spool and housing bore takes the flow direction from the higher to the lower pressure chamber and is largely determined by the gap geometry. The gap width l is defined as the axial overlap of the spool and the bore. The clearance δ is defined as half the diameter difference of the spool and the bore in (7). The relative eccentricity ε is the ratio of the absolute eccentricity e and clearance δ, defined in (8). [12], [13]

$$\delta = \frac{D_{bore} - D_{spool}}{2} \tag{5}$$

$$\varepsilon = \frac{e}{\delta} \tag{6}$$

Under laminar, isothermal flow conditions and a large ratio of gap width l to gap height δ (0.001–0.02 mm) the leakage flow $q_{leakage,max}$ maximizes with the maximum of the eccentricity (ε=1) to 2.5 times the concentric arrangement and approximately follows the equation (7).

$$q_{leakage,max} = \frac{\Delta p \times \pi \times d_m \times \delta^3}{12 \times \eta \times l} \times (1 + 1.5 \times \varepsilon^2) \tag{7}$$

The quantitative calculation of the exact leakage flow is significantly dependent on the real-time viscosity of the fluid at the time of measurement. The adiabatic change of state cannot be assumed since the mass flow rate of the fluid is minor compared to the mass of the solids, so that the fluid temperature is supposed to converge with the temperature of the solids. Additionally, the different coefficients of thermal expansion sensitively affect the annular gap dimension during operation. Consequently, the temperature level of the environment and the connected components has an impact on the leakage measurement. [12], [13]

4. Meta-Regression Model

The underlying problem is to differentiate the impact caused by the gap geometry and by the thermal conditions of the valve in the test bench. A multilinear dependence of the variables on the temperature level is assumed. The goal of designing a meta-regression model is a hand-crafted feature that represents the temperature level of the technical system based on a given range of clearances δ expected to appear mostly in the production dataset and time variables. This approach is intended to eliminate or at least reduce the predictive influence of the clearance on the internal leakage flow, as the clearance is not included in the production dataset.

The test process has a temperature dependent impact. The directional valve is installed between two plates of the testing bench, which is implemented as a rotary indexing table with four workstations. The test program performs two relevant sets of automated testing steps at two successive workstations. The first relevant workstation is characterized by testing steps with high volume flows thus introducing high energy impact into the test part, the mounting plates and pipes. In contrast, testing steps with low volume flows are performed at the second relevant workstation including the label measurement under consideration in the neutral position. At both relevant workstations, a repetition or partial run of a test set is possible, resulting in different temperature scenarios.

4.1 Meta-Training Dataset

The meta-training dataset $\mathcal{D}$ of the directional control valve is provided by Bosch Rexroth AG for this investigation and consists of the shape and geometry data of the components, the pairing data and the metric sensor values measured on the test equipment of the series production. The information from about 20 sensor signals ranges from pressure, current, volume flow and position at around 70 different test steps. In sum, this results in approximately 1000 dimensions per test series. The following procedure is performed for 16 directional control valves. The sample size is limited to 16 at the time of the study due to the significant effort required for geometric measurement, manual bench assembly, and off-series hydraulic testing. Due to varying test durations, the target measurement is recorded with varying frequency for the same temperature scenarios, resulting in a total of 248 test series. In total, a data matrix with 248 rows of 1000 columns is used as the meta-training dataset. Due to a non-disclosure agreement, the dataset cannot be made available to the public.

To represent warm-up and cool-down phases in the system, two time-variables are inserted into the meta-training and production datasets. The warm-up phase is introduced as the runtime sum of the first functional test (FP) workstation with high energy load, while the cool-down phase is defined as the transition time between the end of the first and the beginning of the second relevant workstation with low energy load, including the rest time. In Table 1, the design of the experiment is described by the default setting and variation of the considered variables. The default setting of performing the first set of test steps only once is changed to two and three times to cover different temperature scenarios. The leakage measurement is performed several times for the same heating scenario, as the test part cools down during the measurement.

Table 1: Experiment structure

variable	standard setting	variation
clearance	tolerance centre	± 1.5 µm
number of testing runs	1	1, 2, 3
transition_time	23 sec	number of testing runs
fp_runtime_sum	continuous	number of testing runs

4.2 Meta-learning for Mixed Linear Regression

Following the recommendation of KONG ET AL., a meta-learning for mixed linear regression is developed with the time variables, the ring gap, and the resulting t_{factor} formulated in (8) [10]. The meta-dataset is split into a train and a test split with a split ratio of 85/15. In addition, a five-fold cross-validation is performed to obtain more reliable regression results. To apply this to a sample size of 16, the train dataset must contain 14 valves, while the validation and unseen test datasets must each contain one valve. The selection of valves for each dataset predetermines the result for the coefficients. Therefore, different combinations are iterated for the choice. To increase the repeatability of the results, an equal distribution of the target variables for the training, validation and test datasets is required. The Kolmogorov-Smirnov two-sample test checks for equal distribution for the datasets [14], [15]. Only 42 of the 120 possible combinations pass the Kolmogorov-Smirnov two-sample test with a p-value of 5% and thus only those 42 are considered.

$$t_{factor} = c_1 \times t_{transition} + c_2 \times t_{fp_{runtime}} + c_3 \times \delta^n \quad (8)$$

The median is applied to the probability density of the linear coefficients of the meta-regression model for exponent n = 1 and plotted as a vertical line in Figure 3 to include the most common ratio of c1/c2 for a given annular gap in the production dataset. Considering the heat direction, the transition time for heat dissipation consequently receives a negative sign and the runtime sum for heat input receives a positive sign.

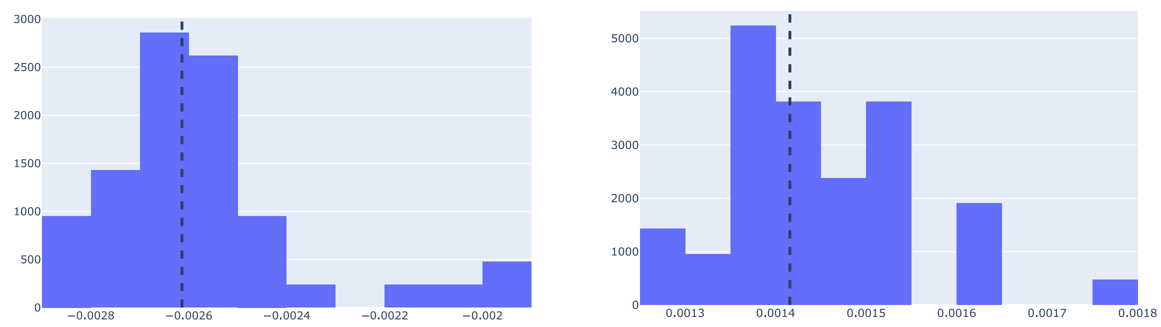


Figure 3: probability density of the linear coefficients

Since the clearance enters the leakage flow formula (7) polynomially and has the highest relevance, the medians from the distributions of the coefficients for different exponents of the clearance are examined and presented in Table 2.

Table 2: Overview of linear and polynomial regression coefficients

	linear	polynomial	
n (clearance)	1	3	50
$t_{transition}$	**-0.002614**	**-0.002580**	**-0,002458**
$t_{fp\_runtime\_sum}$	0.001416	0,001420	0,001182
δ	6.5	0,026222	2,5930

From the overview, it can be concluded that the polynomial regression approach on this meta-training dataset does not produce a significant change in the regression coefficients for the time variables. Since the Gini importance for the transition time in the production dataset is significantly higher compared for the runtime sum, and the regression coefficient for the transition time remains nearly constant, polynomial regression does not produce an improvement in this case.

Since clearance does not occur in the manufacturing dataset, the ratio of the time variables is kept in the hand-crafted feature and follows formula (9) with the regression coefficients of the linear regression.

$$t_{factor} = -0.002614 \times t_{transition} + 0.001416 \times t_{fp_{runtime}} \quad (9)$$

5. Performance Evaluation with Real Manufacturing Data

The prediction model for forecasting leakage values based on series manufacturing data does not contain any geometry data for the bore under consideration. The geometric pairing of the components differs in the manufacturing tolerances of the housing bore and the control supplied. The daily deviation of the internal leakage measurement is caused, among other things, by systematic batch effects of the machining process. The pairing of spool and main housing bore is not cost-effective, so the internal leakage flow is checked as process validation.

First, the manufacturing dataset and the baseline comparison model (A) are described. The enhanced model also includes the time variables (B) and additionally the hand-crafted feature (C). The models are developed in an Anaconda environment using Python version 3.7.9 and XGBoost package version 1.2.1.

5.1 Manufacturing Dataset

With the exception of the geometry, the manufacturing data have the same structure as those of the meta-training dataset. The data for this product was created over the course of six months as part of a product development project. Accordingly, there are slight changes in the process sequence as well as practical inconveniences such as sensor jumps in the period under consideration. For this reason, a thorough cleaning of the data is necessary, during which approximately 3800 data series with about 1000 data columns can be analyzed. Within this dataset, there are numerous retests of the two test sequences so that different temperature levels occur in the system. In addition, tests were performed at all times of the day and different times of the year by different workers.

5.2 Baseline Model

For the baseline XGBoost model, a time series split with a split ratio of 80/20 and a 5-fold cross-validation is employed. A simple grid search is applied for the selection of hyperparameters and named in parentheses below. Hyperparameters not mentioned are set to default. XGBoost squared loss regression (reg:squarederror) is used as the optimization function. The subsample ratio for each boosting iteration was set to 0.8, and the learning rate was set to 0.05 to avoid overfitting. For the same reason, the default of the maximum depth of the trees is lowered from 6 to 4 (max_depth=4) and an early stopping of the training was introduced when the optimization metric does not improve after 15 rounds (early_stopping_rounds=15). Another way to make the model more conservative is to increase the ridge regularization (lambda=20) and the minimum loss reduction for another partition at a leaf node to 100 (gamma=100). Additionally, the minimum sum of instance weight needed in a child is set to 5 (min_child_weight=5). A random seed is implemented to allow comparison of the models.

5.3 Enhanced Model and Final Evaluation

For the enhanced XGBoost model, the baseline XGBoost model is first augmented with the time variables and then additionally with the hand-crafted feature so that the enhancement due to the additional information in the model can be separated from the enhancement due to feature engineering.

The performance of the predictive models using manufacturing data is evaluated using the following standard metrics: root mean square error (RMSE), mean absolute error (MAE), the mean absolute percentage error (MAPE), and coefficient of determination (R^2). Table 3 provides an overview of the metrics for each of the training and unseen test dataset for all three models. In this use case, the MAE and R^2 are most useful for interpretation.

When analyzing the performance evaluation from Table 3, it can be noted that, surprisingly, the metrics of Model B deteriorate in the training dataset compared to the baseline model, but improve slightly in the test dataset. All metrics improve from the baseline model to the extended model C in both the training and test datasets, which plead for the approach to hand-crafted feature engineering based on physical ideas.

Table 3:Performance evaluation

Model name	Model description	train-dataset metrics				test-dataset metrics			
		rmse-mean	mae-mean	mape-mean	r^2-mean	rmse-mean	mae-mean	mape-mean	r^2-mean
A	baseline-model	2.9432	**2.3123**	0.1345	**0.5643**	3.8232	**3.0176**	0.1758	**0.2621**
B	A + time variables	2.9558	2.3215	0.1346	0.5605	3.8104	3.0070	0.1747	0.2669
C	B + t_factor	2.8538	**2.2443**	0.1304	**0.5904**	3.8022	**2.9979**	0.1744	**0.2702**

Regardless of feature engineering, there is a large generalization gap across models between the training and test datasets, suggesting an information gap in each of the manufacturing and meta-training datasets. There are two perspectives to look at the change in metrics: First, the improvement is predominantly in the second decimal place, which certainly does not justify the effort to do so. Second, it is worth noting that the prediction can be improved with a ratio of 16 experiment components to over 3500 production components and despite the existence of a huge information gap in the datasets. The improvement in all metrics is not a coincidence and is a starting point to push this approach further.

Finally, we will mention three justified assumptions for missing information in the dataset. First, the variation of clearance in the production dataset was assumed to be ± 1.5 µm for the meta-training dataset. This range is assumed to be very small and needs to be increased. Second, the sample size for the meta-training dataset is too small compared to the size of the production dataset, as this underrepresents the influence with the greatest predictive power. Third, the time variables used to represent temperature levels are less substantial compared to direct temperature measurement of the components and fluid.

6. Conclusions

The results of the leakage volume flow measurement are very sensitive to changes in gap geometry and temperature level in the measurement cross-section. Using a meta-regression model, a hand-crafted feature based on thermal behavior assumptions can be derived that improves all metrics for predicting leakage volume flow in the post-commission range despite a large cross-model generalization gap between the training and test datasets.

7. Future Work

Referring to the assumptions on the information gap in the datasets, we will significantly increase the sample size for the meta-dataset with a larger span of clearances. Moreover, experiments will follow from which the saturation curves for the time variables can be derived. In addition to more precisely defined time variables, the number of thermal features will be increased by installing more direct temperature measurements in the test bench, if possible, and including them in the manufacturing dataset.

Furthermore, the minimum data acquisition for sufficient prediction accuracy will be investigated. Finally, the quality prediction will be implemented in the serial production of hydraulic valves to predict leakage test results and contribute to a more efficient and competitive manufacturing.

Acknowledgements

The authors would like to additionally thank Dr.-Ing. Matthias Möller and Bosch Rexroth AG for their support and access to their real-world data.

References

[1] Posa, A.; Oresta, P., Lipplois, A.; 2012. Analysis of a directional hydraulic valve by a Direct Numerical Simulation using an immersed-boundary method, Energy Conversion and Management 65, pp. 497–506.

[2] Lisowski, E.; Czyzycki, W.; Rajda, J.; 2013. Three dimensional CFD analysis and experimental test of flow force acting on the spool of solenoid operated directional control valve, Energy Conversion and Management Volume 70, pp. 220–229.

[3] Weichert, D.; Link, P.; Stoll, A.; Rüping, S.; Ihlenfeldt, S.; Wrobel, S.; 2019. A review of machine learning for the optimization of production processes, Springer, London.

[4] Köksal, G.; Batmaz I.; Testik MC; 2011. A review of data mining applications for quality improvement in manufacturing industry. Expert System Applications 38 (10), pp. 13, 448–13, 467.

[5] Lisowski, E.; Rajda, J.; 2013. CFD analysis of pressure loss during flow by hydraulic directional control valve constructed from logic valves, Energy Conversion and Management 65, pp. 285–291.

[6] Wuest, T.; Klein; D.; Thoben; K.-D.; 2011. State of steel products in industrial production processes, Procedia Engineering 10, pp. 2220–2225.

[7] Wuest, T.; Irgens, C.; 2013. An approach to monitoring quality in manufacturing using supervised machine learning on product state data, Springer Science+Business, New York.

[8] Duboue, P.; 2020. The Art of Feature Engineering. Essentials for Machine Learning, pp. 79, 124–126, Cambridge University Press.

[9] Cai, D.; Sheth, R.; Mackey, L.; Fusi, N.; 2020. Weighted Meta-Learning, arXiv:2003.09465v1.

[10] Kong, W.; Somani, R.; Song, Z.; Kakade, S.; Oh, S.; 2020. Meta-learning for mixed linear regression, Proceedings of the 37th International Conference on Machine Learning.

[11] Lovrec, D.; Kastrevc, M.; Ulaga, S.; 2008. Electro-hydraulic load sensing with a speed-controlled hydraulic supply system on forming-machines, pp. 474, 537–543, Springer Vieweg, London.

[12] Findeisen, D.; Helduser, S.; 2015. Ölhydraulik. Handbuch der hydraulischen Antriebe und Steuerungen, pp. 66–68, 474, 537–543, Springer Vieweg, Berlin.

[13] Matthies, H.J.; Renius, K.T.; 2014. Physikalische Grundlagen ölhydraulischer Systeme, in: Einführung in die Ölhydraulik, pp. 48–50, Springer Vieweg, Wiesbaden.

[14] Hodges; J.; 1958. The significance probability of the smirnov two-sample test. Arkiv för Matematik, 3, No. 43, pp. 469–486.

[15] Lopes, R; Reid, I.; Hobson, P.; 2007. The two-dimensional Kolmogorov-Smirnov test. Proceedings of Science.

[16] Wang, J.; Ma, Y.; Zhang, L.; Gao, R.; Wu, D.; 2018. Deep learning for smart manufacturing: Methods and applications, Journal for Manufacturing Systems 48, pp. 144–156.

[17] Wang, Y.-X.; Ramanan, D.; Hebert, M.; 2017. Learning to model the tail; In: Advances in Neural Information Processing Systems, pp. 7029–7039.

[18] Li, Z.; Zhou, F.; Chen, F.; Li, H.; 2017. Meta-sgd: Learning to learn quickly for few-shot learning. arXiv preprint arXiv:1707.09835.

[19] Bertinetto, L., Henriques, J. F., Torr, P. H.; Vedaldi, A.; 2019. Meta-learning with differentiable closed-form solvers. International Conference on Learning Representations. arXiv:1805.08136, 2019.

[20] Zhou, F.; Wu, B.; and Li, Z.; 2018. Deep meta-learning: Learning to learn in the concept space, arXiv preprint arXiv:1802.03596.

[21] Zintgraf, L.; Shiarli, K.; Kurin, V.; Hofmann, K.; White-son, S.; 2019. Fast context adaptation via meta-learning. In: International Conference on Machine Learning, pp. 7693–7702.

[22] Rajeswaran, A.; Finn, C.; Kakade, S. M.; Levine, S.; 2019. Meta-learning with implicit gradients. In: Advances in Neural Information Processing Systems, pp.113–124.

[23] Grant, E.; Finn, C.; Levine, S.; Darrell, T.; Griffiths, T.; 2018. Recasting gradient-based meta-learning as hierarchical bayes, arXiv preprint arXiv:1801.08930.

[24] Finn, C.; Xu, K.; Levine, S.; 2018. Probabilistic model-agnostic meta-learning. In: Advances in Neural Information Processing Systems, pp. 9516–9527.

[25] Kim, T.; Yoon, J.; Dia, O.; Kim, S.; Bengio, Y.; Ahn, S.; 2018. Bayesian model-agnostic meta-learning. 32nd Conference on Neural Information Processing Systems, arXiv preprint arXiv:1806.03836.

[26] Harrison, J.; Sharma, A.; Pavone M.; 2020. Meta-learning Priors for Efficient Online Bayesian Regression. In: Morales M., Tapia L., Sánchez-Ante G., Hutchinson S. (eds) Algorithmic Foundations of Robotics XIII. WAFR 2018. Springer Proceedings in Advanced Robotics, vol 14. Springer, pp. 318–337.

[27] Chen, S.; Li, J.; Song, Z.; 2020. Learning mixtures of linear regressions in subexponential time via Fourier moments. Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing, pp. 587–600.

[28] Friedman, J.; 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics, 29 (5); pp. 1189–1232.

[29] Kiangala, S.K.; Wang, Z.; 2021. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment, Machine Learning with Applications, 4.

[30] Chen, T.; Guestrin, C.; 2016. XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Biography

Christian Neunzig has been a doctoral student 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 real production data. The dissertation is supervised by Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter.

Simon Fahle has been a member of the Chair of Production Systems (LPS) at Ruhr-Universität Bochum since 2019. He completed a bachelor's and master's degree in mechanical engineering at Ruhr-Universität Bochum. His primary research topics are machine learning, time series and radial-axial ring rolling.

Bernd Kuhlenkötter was responsible for product management and engineering 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. Since 2015, Univ. Prof. Dr.-Ing. Bernd Kuhlenkötter has held the professorship for "Production Systems" at the Ruhr University Bochum.

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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Quantification of a System Dynamics Model for Optimized Failure Management in Manual Assembly

Junjie Liang[1], Robin Guenther[1], Sebastian Beckschulte[1], Robin Exner[2]
Raphael Kiesel[1], Robert H. Schmitt[1, 3]

[1]*Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Aachen, Germany*
[2]*Kostal Industrie Elektrik GmbH, Hagen, Germany*
[3]*Fraunhofer Institute for Production Technology IPT, Aachen, Germany*

Abstract

Companies have only limited resources to carry out value-adding activities. In order to achieve high quality requirements, the goal-oriented and efficient handling of failure incidents is a competence that must be emphasized. However, the activities associated with this represent an additional effort that consumes part of the resources available. For this reason, it is important to optimally coordinate the value-adding processes of service creation on the one hand and the processes of failure management on the other. Accordingly, the objective of this paper is to program a simulation model for an optimized failure management in manual assembly, which transforms the failure management from traditionally experience-based to model-based. To achieve this objective, the consideration of the interactions between the failure management process and operational activities during the production process is essential. However, in current literature, interactions between production and failure management still lack detailed descriptions. Thus, both disciplines are often considered and optimized in isolation. Therefore, an advanced System Dynamics Model representing manual assembly processes with 23 elements involved is constructed and applied to indicate the interactions between production and failure management. This enables the optimized configuration of the failure management activities depending on the circumstances to be taken into account. According to the generated model, a generic process module is programmed and test runs are performed to assess the model behaviour's plausibility. The programmed System Dynamics model is implemented and validated in a use case of a manual assembly line consisting of two assembly stations. For this purpose, the generated process model is linked to the production chain of the use case and parameterized accordingly. This procedure demonstrates that the model can be used to derive general recommendations for action in order to realize an optimized design of failure management activities.

Keywords

Failure Management; Failure Handling; Manual Assembly; System Dynamics

1. Introduction

Manufacturing companies face the constant challenge of increasing quality requirements, shortened product life cycles and cost-efficient production. A major force influencing the achievement of quality, time and cost targets is the occurrence and respective handling of failures. Reference processes for the long-term elimination of failures are widely established and single steps for handling a failure are well known [1]. One problem with these processes is that they are usually designed without taking into account the interactions with the actual value adding processes. Despite the close connection between these two domains, failure management activities are often defined and optimized in isolation [2]. Thus, it is not considered which positive or negative consequences a failure management process can have for the company as a whole. This circumstance is to be considered critically, since a conflict of objectives can exist between medium to long-term oriented failure management activities and short-term oriented production objectives with regard to the use of available resources [3]. Consequently, in order to optimize failure management, the reciprocal influence of failure prevention and product manufacturing must be taken into account. According to these thoughts, the development of a simulation model for the derivation of recommendations is described in the following. The objective is the effective and efficient integration of the failure management process into the production regarding a use case of manual assembly. Here, the hypothesis is followed that recommendations for action can be derived on the basis of a System Dynamics model. Using simulation analyses of a real assembly chain, recommendations for optimizing failure management, i.e., measures for better networking of a company's failure management strategy with its production planning and control, are derived taking into account company-specific factors.

2. State of the art

Failure management is an essential topic for manufacturing companies. Reduction of diverse types of failures during production is the critical key to achieve a better performance of a manufacturing company [4]. The concept of failure management in the manufacturing domain includes all measures for the sustainable avoidance and prevention of failures to ensure the success of production [5]. Concerning the timing of activities along the production value chain, the existing approaches for failure management can be divided into preventive methods and reactive methods. Preventive methods aim at the prevention of failure before occurrence. One of the most well-known methods is FMEA (Failure Mode and Effects Analysis) [6]. It is applicable for product design and product manufacturing, and it aims to anticipate possible failures during the manufacturing of the product under consideration. In contrast, reactive failure management is only used when failure avoidance is no longer possible, and the failure has already occurred [7]. There are various methods that describe how to deal with these failures. As we mainly aim at a model-based management of occurred failures in production domain – particularly manual assembly – and enabling manufacturing companies to learn from their existing failures, reactive failure management methods are of high importance and will be considered in the following analysis. Simulation, data modeling and many other approaches already exist in order to make the internal failure management during production process as effective and efficient as possible. Various approaches for reactive failure management are introduced by different authors.

ELLOUZE developed SAFE to support producing companies in applying reactive measures to eliminate failures that have already occurred [8]. This failure management approach includes methodological support for workers. The aim is to transmit the necessary information to the right worker at the right time in the right place. The core of this approach is to establish comprehensive and systematic storage of failure knowledge. The knowledge can be forwarded to the worker through integration into Workflow-Management-System. The failure occurrences lead to increasing gained failure knowledge. Based on this constantly growing knowledge, the decision quality concerning the failure priority and the selection of measures should be continuously increased.

The project FAMOS dealt in particular with the requirement-based input of failure information and the finding of existing knowledge within the failure knowledge database [9]. In this approach, ICT (information and communications technology) stands for the storage and use of failure knowledge within the producing company. The method assists workers in correlating suspected failure patterns to database items for which actions have previously been established. Workshops are explicitly initiated for the identified failure patterns, and measures/actions are developed to store them in the database subsequently. Thus, the approach from

FAMOS does not provide a reference process for failure management but comprises a set of IT-supported methods for creating a robust failure knowledge database.

Different from the approach from FAMOS, LINß described a reference process for failure elimination [5]. He defined seven necessary process steps, including production failure definition, failure recording, immediate action, failure analysis, corrective measures development, corrective measures implementation, and results controlling. LINß distinguished between external (cf. complaint management) and internal (cf. failure management) failure correction. His model describes the necessary tasks during failure elimination.

The approaches listed above are considered isolated failure management approaches. These isolated approaches specifically focus on the failure management processes and distinguish them from other processes in the producing company. Reference processes to be followed for successful failure elimination are mostly mentioned. We consider that providing reference process for optimized failure elimination receives special attention from many authors. What is predominantly not included in the above-mentioned approaches, however, is the consideration of the other processes in companies. Instead, the authors describe a self-sufficient design of failure management without addressing possible conflicts or intersections with other processes. In particular, little attention is paid to the interaction with the value-creating processes. Hence, the integration of failure management into the value adding production process is the focus in our work.

3. Programming of the simulation model

As described in our previous work [10], the analysis of existing failure in a production system is based on the proposed modularized analysing method according to TUERTMANN ET AL. [11]. The model observes the production system in a modularized way. Based on the modules, the construction of the failure management model in our previous work was accordingly derived. The failure management model then includes *Failure Management*, *Failure Knowledge*, *Failure Causes*, *Resources* and *Production*, extracted from the work of TUERTMANN ET AL. It is to be mentioned that in this model no individual process steps are represented, but only the manufacturing company's entire production in one calculation step. Interaction among different steps can thus not be shown.

Generally, in the model of TUERTMANN ET AL., the selected level of abstraction is too high to enable actual optimization at the level of production planning and control. In particular, the level of detail presented within the model does not allow for a simulation at the process step level, as this is necessary for calculating the key parameters to achieve better production performance. The failure management model should be enabled to map any production process in order to be able to consider specific scenarios. Hence, in our previous work [10], adjustments and further development are made to adapt the model from TUERTMANN ET AL. and make it applicable for failure management in manual assembly. A schematic structure is proposed to modularize the production steps and enable representing any number of production steps in a single model. As next step, this model needs to be programmed considering the interaction among the process steps. The corresponding parameters involved in the model need to be described in mathematical forms.

For the programming of the simulation model presented in our previous paper [10], the software *Vensim* developed by *Ventana Systems* is applied. This software was also deployed by TUERTMANN ET AL., whose model serves as the origin of the simulation model in this paper [3]. Deploying the software *Vensim* for programming the simulation model can reduce the risk of transmission failures. The programmed simulation model consists of 23 elements and is not restricted by the amount or the type of the process step. These 23 elements include ten constants, ten variables, and three additional variables with integrating behaviour. The constants are the setting parameters of the simulation model. By setting corresponding values of constants, the model can be adapted to a specific process step (PS). Table 1 provides detailed information about the constants.

Table 1: Constants of the simulation model and their description and unit

Constant	Description	Unit
Acceptable Defect Rate PS	Percentage of defective products accepted without initiating activities to correct the defect	%
Cycle Time PS	Execution time of the considered process step	Seconds/ Product
Defect Appearance Probability PS	Probability with which a present failure cause leads to an actual failure occurrence	%
Defect Registration Rate PS	Proportion of detected defects that is recorded	%
Inspection Frequency PS	Interval of product inspection	Products
Inspection Time PS	Duration of the inspection of one product	Seconds/ Product
Planning PS	Number of planned products for the process step	Products/ Day
Resource Requirement for One Defect Registration PS	Resource requirements for the recording of a defect	Seconds/ Defect
Shift Length PS	Length of a shift excluding breaks	Hours
Shifts per Day PS	Number of shifts per day	Shifts

Aside from setting parameters, variables are also involved during simulation model programming. The variables are subject to mathematical calculation rules, which must be represented in the programmed simulation model. The variables with descriptions and calculation rules can be found in Table 2.

Table 2: Variables of the simulation model and their description and calculation rule

Variable	Description	Calculation Rule
Completion Rate PS	Number of finished products related to one time unit	$\text{Completion Rate PS}=\text{Delay Fixed}(\frac{\text{Net Production Time PS}}{\text{Cycle Time PS}}*3600, \frac{\text{Cycle Time PS}}{86400}, 0)$
Conform Products Rate PS	Product flow of the compliant products	Conform Products Rate PS=Completion Rate PS*(1-Defect Rate PS)
Defect Rate PS	Percentage of defectively produced products in the process step	Defect Rate PS=Defect Appearance Probability PS*Root Causes
Defective Products Rate PS	Difference between all finished products and compliant products	Defective Products Rate PS $=\text{Delay Fixed}(\text{Defect Rate PS} * \text{Completion Rate PS}, \frac{\text{Inspection Time PS}}{86400}, 0)$
Need for Action PS	Need for action resulting from the comparison of the acceptable defect rate and the actual defect rate	Need for Action PS =IF THEN ELSE(Defect Rate PS>Acceptable Defect Rate PS,1, 0)
Net Production Time PS	Working time in hours available per day	Net Production Time PS =Required Worker PS*Shifts per Day PS*Shift Length PS -Time Spent on Inspections PS $-\frac{\text{Time Spent on Defect Registration PS}}{3600}$
Required Worker PS	Number of employees required per day and shift	Required Worker PS =Required Working Time PS/(Shift Length PS*Shifts per Day PS)

Required Working Time PS		Required Working Time PS $=\text{Planning PS}*\frac{\text{Cycle Time PS}}{3600}+\text{Time Spent on Inspections PS}$
Time Spent on Defect Registration PS	Time required for the registration of defects	Time Spent on Defect Registration PS =Defect Registration Rate PS*Defective Products Rate PS *Resource Requirement for One Defect Registration PS
Time Spent on Inspections PS	Time required for quality controls	$\text{Times Spent on Inspections}=\frac{\text{Planning PS*Inspetion Time PS/3600}}{\text{Inspection Frequency PS}}$

Different from the ten variables above, variables with integrating behaviour are primarily intended to model the material flow of the process step. The initial value of these variables equals zero. The descriptions and calculation rules of the variables with integrating behaviour are given in Table 3.

Table 3: Variables with integrating behaviour of the simulation model and their description and calculation rules

Variable	Description	Unit
Completed Products PS	Completed products of the process step	Completed Products PS =Integ(Completion Rate PS-Conform Products Rate PS-Defective Products Rate PS)
Conform Products PS	Conform products of the process step	Conform Products PS=Integ(Confrom Products Rate PS)
WIP Process Step	Work occurring in the process step	WIP Process Step=Integ(Defective Products Rate PS+Planning PS-Completion Rate PS)

In addition to the simulation model itself, the target system for performance measurement must be integrated. Only by measuring performance, the evaluation of the various failure management strategies is enabled. Performance refers to the time required for an employee to process a conform product. The overall target value of one process step is calculated as follows:

$$\text{Overall Target Value PS=IF THEN ELSE(Conform Products Rate PS>0, }\frac{\text{Required Working Time PS}}{\text{Conform Products Rate PS}}\text{,0)} \quad (1)$$

To compare two simulation results, a further variable with integrating behaviour (Sum of OTV PS) is added. Thus, the performance achievement over the entire simulation period can be compared in only one characteristic value. The initial value is zero and the calculation rule for the variable is:

$$\text{Sum of OTV PS=Integ(Overall Target Value PS)} \quad (2)$$

3.1 Verification of the programmed simulation model

Towards the error-free programming process, tests were already carried out during model generation. It was checked after each programmed calculation rule using dimension tests whether the units used for the elements were consistent. Additionally, to check the programmed simulation model's behaviour for plausibility and to ensure that the generated links of the model elements follow logical relationships, test runs were performed. The procedure is based on the techniques of verification and validation, according to RABE ET AL. [12]. Such a test is considered a fixed or limiting value test. In this test, the simulation model is given fixed input parameters to prevent dynamic model behaviour; hence, the simulation results can be reproduced. Accordingly, fictive values were assigned to the simulation model's constants, and hypotheses were made for the resulting values of the variables. The test run confirmed all hypotheses so that a consistently alleged behaviour of the calculation rules and implemented functions are proven.

Furthermore, it was tested whether the system behaviour remains stable when the input values are changed. For this purpose, several adjustments are made to the initial model, and the expected change in system behaviour is compared with the simulation result of the adjusted model. In each case, the initial model forms the basis for the adaptation, i.e., after each adaptation, the model is reset to the initial state before the subsequent adaptation is carried out. The constants *Defect Appearance Probability PS*, *Planning PS*, and *Cycle Time PS* were respectively adjusted. The expected results after adjusting the first and second constants were confirmed. The expected result after adjusting the cycle time deviates 0.16 % from the actual result. However, the deviation decreases when the time spans between the calculation times are reduced.

3.2 Validation of the programmed simulation model using a case study

The programmed simulation model was validated using a case study from the automotive industry. The validation of the concept is carried out by comparing the simulation results with the real recorded data. Two successive assembly stations for the production of a powertrain module are considered. For each station, one employee works autonomously to carry out all the tasks required at the station. If several workers are involved, they perform the same activities in parallel on the same type of product. In the following, the work

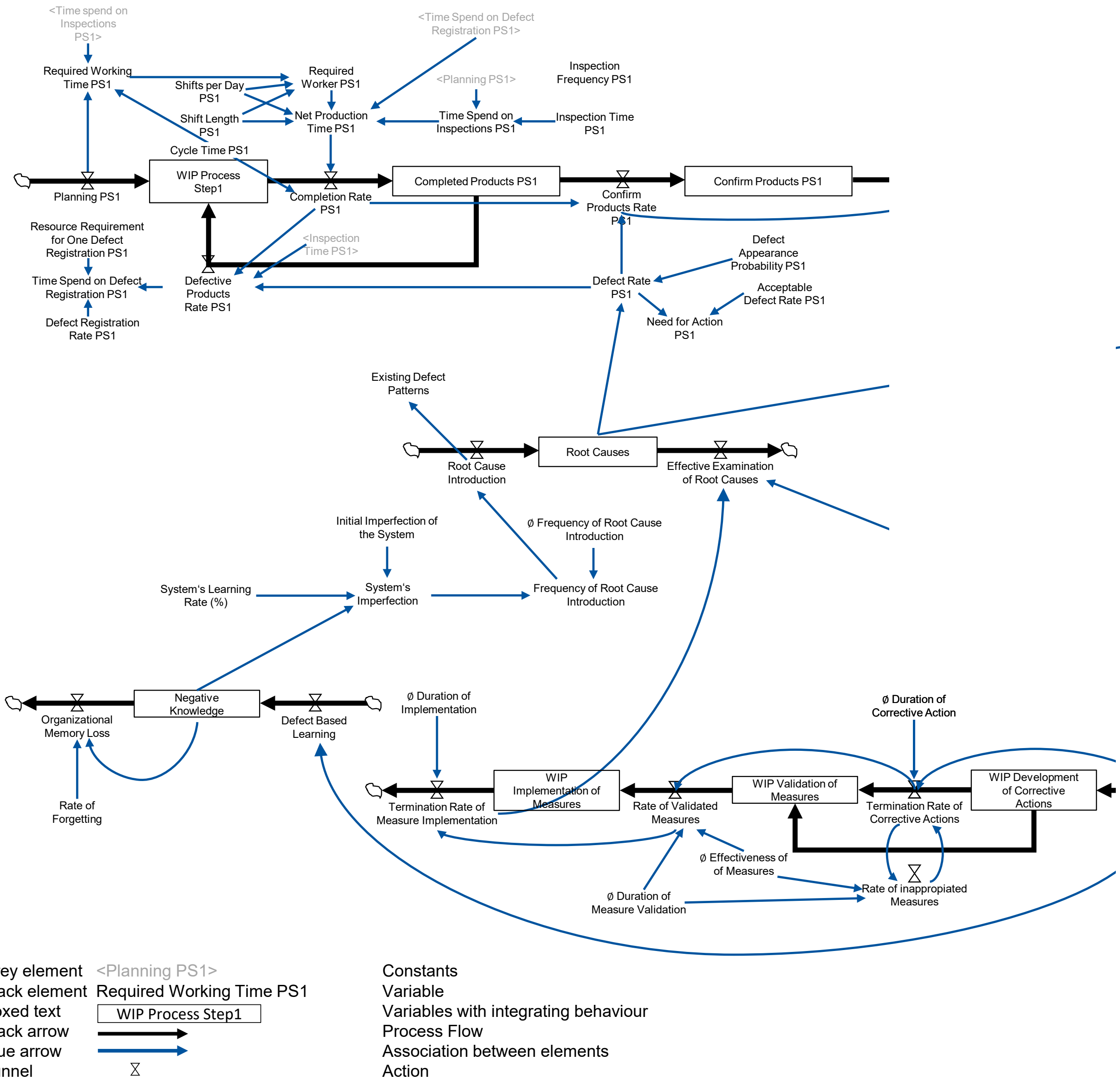


Figure 1: First part of the overall model of the use case

stations are referred to as Station 1 and Station 2 according to the sequence in which they are to be performed. A total of eleven different tasks are performed at Station 1 and eight at Station 2.

To begin the simulation model must be adapted to the case study. Since two process steps are considered, two model modules must be connected in series. It is necessary to link these models with the sub-models Failure Management, Failure Causes, and Resources. These sub-models can be taken over to a large extent from the approaches according to TUERTMANN ET AL. [3]. An adaptation of the sub-models must be carried out primarily regarding the respective interfaces. The need for action in failure management is determined based on the need for action in the individual process steps and can no longer follow directly from a model of the entire production process. In Figure 1 and **Error! Reference source not found.**, the overall model of the use case is visualized.

As a sequent step, the failure occurrence probability of stations 1 and 2 must be determined. The probability is determined according to the Expert System for Task Taxonomy (ESAT) [13]. The prerequisite is the

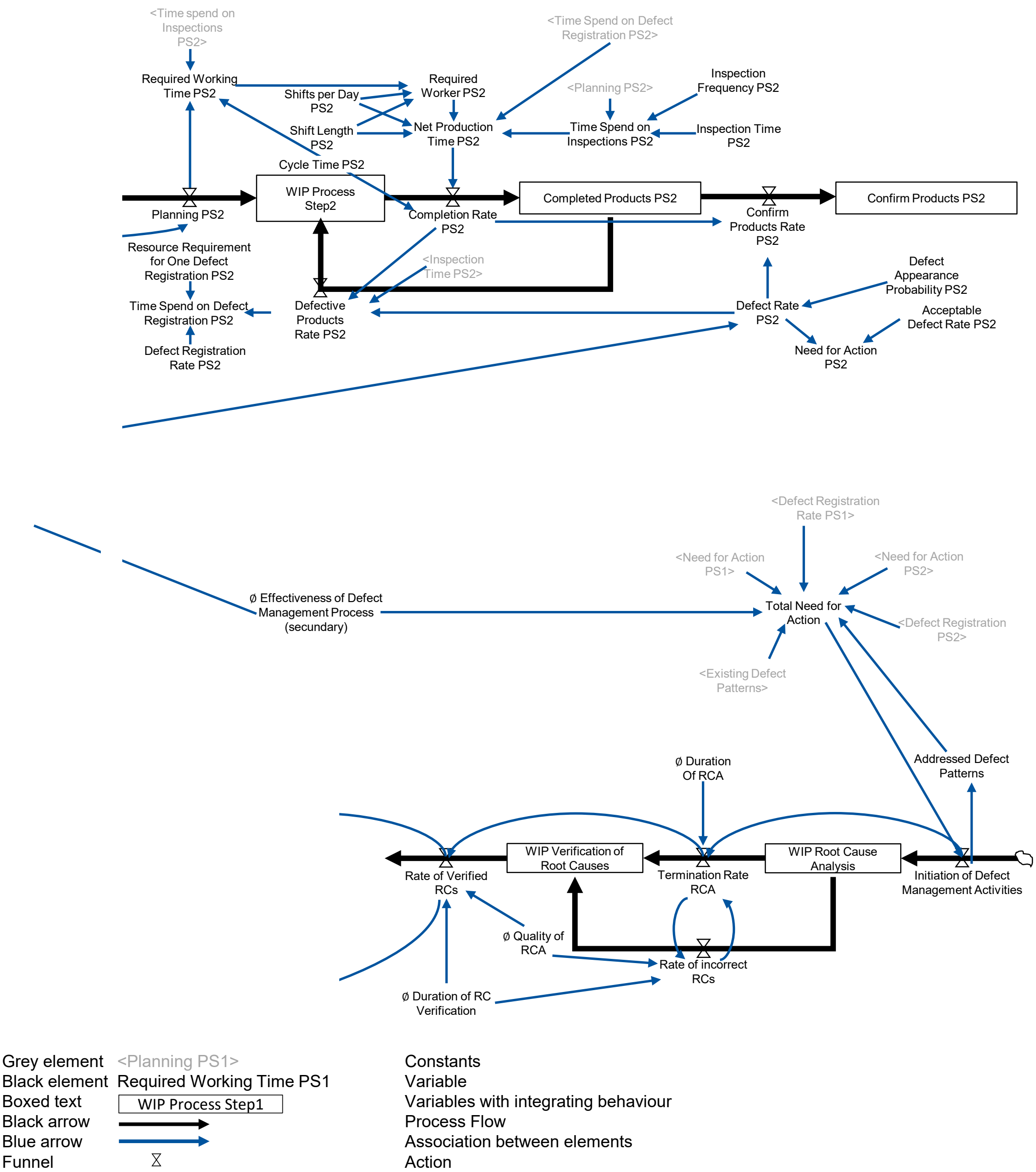


Figure 2: Second part of the overall model of the use case

description of the tasks at stations 1 and 2. A corresponding task catalogue was created in a workshop with the foremen and workers of the assembly line.

Once all other input data, such as cycle time, has been collected for both stations, the simulation model can be executed. In order to calculate the actual failure rates, all units produced in 2016 were considered. In total, 100,000 units were produced at the respective stations during this period. To determine the failure rate, all failures that occurred during this period were considered, and those failure types were selected that could be traced back to a failure at assembly stations 1 or 2. At station 1, the simulation result is 12 % below the actual value, and at assembly station 2 is 1 % above. Assuming that the standard deviation of the absolute number of failures is in each case only one failure (related to one year), a tolerance range in terms of process dispersion can be defined to +/- 3 sigma (station 1: +/- 20%; station 2: +/- 2%). The deviations of the simulation result from the actual value are within this 3-sigma range and are therefore considered as an acceptable deviation. In conclusion, the programmed model is valid [11].

4. Derivation of general recommendations for action for failure handling using a case study

Subsequently, the verified and validated model was used to generate recommendations for the optimal integration of failure management into the production process. For this purpose, another case study was considered, which has a significantly lower process level than the validation case study. This case study also comprises two process steps of a manual assembly line. Simulations were performed over 50 days. In these simulations, the failure management strategy was continuously varied. A failure management strategy is defined as setting values for the constants *Acceptable Defect Rate PS* and *Defect Registration Rate PS*. Accordingly, a variation of the strategy is equivalent to a variation of the input values of these two constants. In order to determine the performance over the entire simulation period, the variable *Sum of OTV PS* was evaluated by relating the result of one run to the result of the baseline situation.

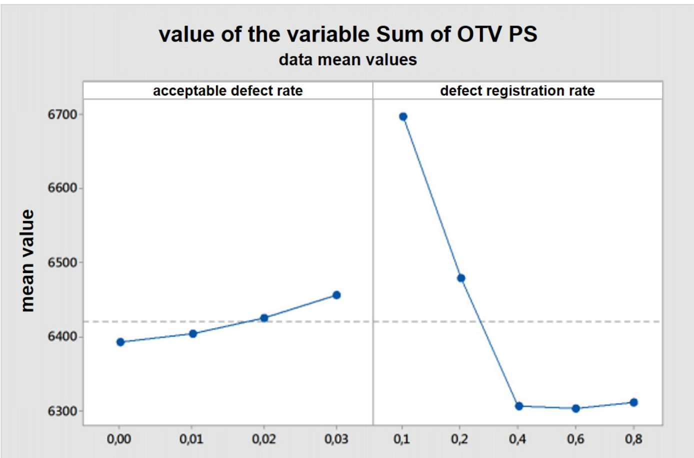


Figure 3: Influence of the failure management strategy on the performance

The best results were obtained for a minimum acceptable defect rate and a defect registration rate around 50%. In addition to the overall strategy's performance, it was also possible to analyse the influences of the acceptable defect rate and defect registration rate on the performance. For this purpose, the defect registration rate was varied at a constant acceptable defect rate. Only after 15 days of simulation and variation, respectively increase of the defect registration rate, an improvement of the performance became visible by the value of the variable *Sum of OTV PS.* From a defect registration rate of approximately 40 % onwards, no further improvement could be identified. According to the simulation results, it is therefore not expedient to capture all failures since after a limit value has been exceeded, no further improvement in performance can be achieved. That indicates that there is an optimal defect registration rate that must be set to maximize performance.

As following step, the acceptable defect rate was varied at a constant defect registration rate. The simulation results show that the influence on the acceptable defect rate's performance is significantly lower than the one of the defect registration rate. The variable *Sum of OTV PS* has its highest value on day 15, and with lowering the rate, the value drops again. To further characterize the influence of the acceptable defect rate and a defect registration rate, variance analysis was performed. For this purpose, the variable *Sum of OTV PS* was defined as the target variable and the two constants are set as factors. In Figure 3, the variance analysis results to characterize the influence of the failure management strategy on the performance are illustrated.

5. Conclusion

Based on the simulation results, it was shown that by adjusting the failure management strategy, the time required for an employee to assemble a compliant product could be reduced. For this adjustment, the defect registration rate and acceptable defect rate are the levers. It was also found that an optimum for the defect registration rate exists. The defect capture rate describes the proportion of existing product defects fed into a failure management process to eliminate the cause of the failure. A defect registration rate below the optimum can lead to a significant drop in performance in the entire production system. Maximization of the defect registration rate towards 100 % causes inefficiency since the resulting additional effort for defect detection can no longer be compensated by the increased performance of failure management.

The results show that a System Dynamics model can be used to derive recommendations for optimizing failure management. This allows measures to be identified for better interconnection of a company's failure management strategy with its production planning and control, taking into account the mutual interactions. However, the efforts of the remaining personnel in failure management are not included in the optimization. There is potential for expansion in terms of integrating a role model so that the efforts of the failure management in the different groups of persons can be included.

The simulation model is based on a static strategy, which means that the defect registration rate and acceptable defect rate are defined initially and are not further adjusted afterward. Especially concerning the consideration of a production ramp-up, an adjustment of the failure management strategy depending on the time seems to be reasonable. The integration of a dynamic strategy is a potential extension possibility of the model to increase the model quality.

Acknowledgements

The support of the German National Science Foundation (Deutsche Forschungsgemeinschaft DFG) through the funding of the research project “Modellbasierte Optimierung der Fehlerabstellung in Produktionssystemen” (GZ: SCHM 1856/71-3 AOBJ: 661218) is gratefully acknowledged.

References

[1] Beckschulte, S., Günther, R., Kiesel, R., Schmitt, R. H., 2021. Quality Improvement through Data Analysis – Qualification of Failure Management by Standardized Failure Recording in Manual Assembly. Proceedings of the 11th Congress of the German Academic Association for Production Technology (WGP), in print.

[2] Exner, R., 2019. Ein quantitatives Modell zur Unterstützung des Fehlermanagements in der manuellen Montage. Apprimus Verlag, Aachen.

[3] Tuertmann, R., Gloeckner, H., Falk, B., Schmitt, R. H., 2017. Conceptual modelling of the failure management process in the manufacturing industry. Total Quality Management & Business Excellence 28 (9), pp. 1041 – 1053.

[4] Hochreither, P., 2005. Fehlermanagement im Unternehmen: wie aus Fehlern Umsatz und Gewinn werden. BusinessVillage GmbH, Goettingen.

[5] Linß, G., 2011. Qualitätsmanagement für Ingenieure, 3rd ed. Carl Hanser Verlag, Munich.

[6] Werdich, M., 2012. FMEA-Einführung und Moderation, 2nd ed. Vieweg+Teubner Verlag, Berlin.

[7] Beaujean, P., 2011. Modular gestaltetes reaktives Qualitätsmanagement. Apprimus Verlag, Aachen.

[8] Ellouze, W., 2010. Entwicklung eines Modells für ein ganzheitliches Fehlermanagement: ein prozessorientiertes Referenzmodell zum effizienten Fehlermanagement. Shaker Verlag, Aachen.

[9] Plach, A., 2011. Entwicklung einer expertenbasierten Fehlermanagement-Methode. Shaker Verlag, Aachen.

[10] Exner, R., Ngo, Q. H., Liang, J., Ellerich, M., Günther, R., Schmitt, S., Schmitt, R. H., 2020. A Failure Handling Process Model for Failure Management in Manual Assembly. Proceedings of the 2020 IEEE IEEM, pp. 1225 – 1229.

[11] Tuertmann, R., Gloeckner, H., Falk, B., Schmitt, R. H., 2017. Analyzing the Feedback Structure of Production Defect Management. 9th International Working Conference, Total Quality Management - Advanced and Intelligent Approaches, Belgrade.

[12] Rabe, M., Spiekermann, S., Wenzel, S., 2008. Verifikation und Validierung für die Simulation in Produktion und Logistik: Vorgehensmodelle und Techniken. Springer Verlag, Berlin.

[13] Brauser, K., 1992. ESAT - Ein neues Verfahren zur Abschätzung der menschlichen Zuverlässigkeit. DGLR Report.

Biography

Junjie Liang (*1995) studied industrial engineering at RWTH Aachen University. Since 2020, he has been working on his doctorate as a research assistant at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University, focusing on production optimization and production data modelling.

Robin Guenther (*1988) studied industrial engineering at RWTH Aachen University. Since 2018, he has been working on his doctorate as a research assistant at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University, focusing on production optimization and failure management.

Sebastian Beckschulte (*1990) studied industrial engineering at the University of Duisburg-Essen. Since 2019, he has been working on his doctorate as a research assistant at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University, focusing on quality intelligence and failure management.

Robin Exner (*1987) is Director Production Engineering at KOSTAL Industrie Elektrik in Hagen**.** He worked as a Team Leader for at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University and graduated in 2019 with a Ph.D. in the field of Failure Management in Manual Assemblies.

Raphael Kiesel (*1991) is chief engineer at the Laboratory for Machine Tools and Production Engineering of RWTH Aachen University. He graduated in Mechanical Engineering from the University of Wisconsin-Madison in 2016, followed by a degree in Industrial Engineering from RWTH Aachen University in 2017.

Robert H. Schmitt (*1961) has been professor of Chair of Metrology and Quality Management and Member of the Board of Directors at Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen since 2004. He is also Member of the Board of Directors at Fraunhofer Institute of Production Technology (IPT).

2nd Conference on Production Systems and Logistics

Model-based Approach for the Automation and Acceleration of the CE-Conformity Process for Modular Production Systems: Future Requirements and Potentials

Jörg Siegert[1,2], Lukas Krispin[1], Awad Ramez[2], Mohamed El-Shamouty[2], Thilo Schlegel[1,2], Liliana Zarco[1], Florian Roth[1], Oliver Mannuß[2]

[1]*Institute of Industrial Manufacturing and Management IFF, University of Stuttgart, Nobelstr. 12, Stuttgart, 70569, Germany*
[2]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstr. 12, Stuttgart, 70569, Germany*

Abstract

In the ultra-flexible factory, production machines reconfigure themselves ad-hoc depending on incoming production orders. Consequently, the IFF institute has developed a modular cyber-physical system (CPS) that consists of configurable robotic components. This paper is motivated by the need to reduce the effort and time associated with the CE conformity process for manufacturing systems with frequent reconfigurations, thus minimizing non-value adding machine downtime and increasing Overall Equipment Effectiveness (OEE). It presents an approach for an automated assistant system for the required risk assessment process using a database of previous CE conform system configurations and a reasoner that is able to verify the existence of a CE conform configuration that supersedes the hazards and risks of the new configuration. This allows the protective measures listed in the previous CE conform configuration to be reused, as well as a large part of the respective CE documentation.

Keywords

CE conformity process; human robot collaboration; safety; modular machines; reconfigurable manufacturing system; Industry 4.0; information model; cyber-physical systems, Matrix Fusion Factory

1. Introduction

Since the early 1980s the global manufacturing industry in general, and the European industry in particular, has had to cope with lower lot sizes, higher numbers of product variants and shorter product life cycles [1]. In response, new manufacturing concepts such as flexible, reconfigurable and changeable manufacturing systems (FMS, RMS and CMS) have emerged [2–4]. These manufacturing system concepts exploit the modularity of hardware and software, as well as model-based engineering, to rapidly adapt the manufacturing system to new product and production requirements. In recent years, the interconnectivity of manufacturing components, artificial intelligence, and digitalization have extended the concepts towards CPS that are able to semi-autonomously analyze product models, reason about the required production reconfiguration, and even partially reconfigure themselves [5–7]. Industry 4.0 has established itself as the common trademark for such systems. One example of such an Industry 4.0 system is the Matrix Fusion Factory (MFF) depicted in Figure 1. This outlines how temporary modular and mobile hardware and

DOI: https://doi.org/10.15488/11273 publish-Ing.

software assemblies (classically referred to as machines) are composed. Humans and modular machines move and interact in the same workspace. Due to the constant adaptation of the production system to the orders, separated workspaces may not be beneficial and may have a negative impact on value-adding [8–10]. The modular assemblies, for example modular assembly I and II in Figure 1, consist of modules (a, a* to f*) with different functions and capabilities. Thus, the required capabilities can be generated flexibly by combining different modules. However, while autonomy, modularity, and flexibility enable manufacturing systems to be reconfigured and adapted more frequently, the logical conclusion of these enablers results in "previously unknown" manufacturing system configurations. These configurations, for which no CE markings have been issued, are potentially hazardous, considering the workspaces potentially shared by workers and machines. This makes CE marking and the corresponding risk/safety assessment the future bottleneck with regard to the frequent reconfiguration of manufacturing systems [11].

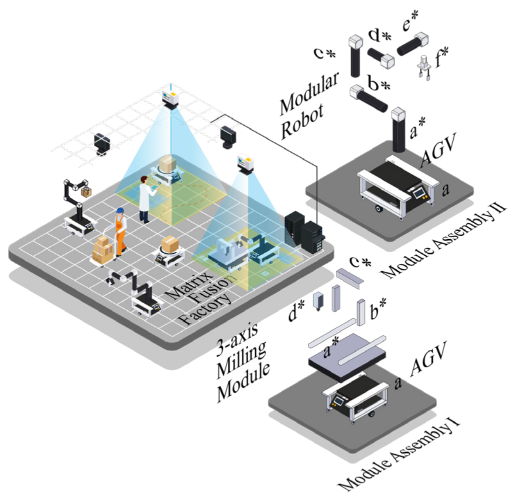


Figure 1: Concept of the Matrix Fusion Factory with modular, mobile machines

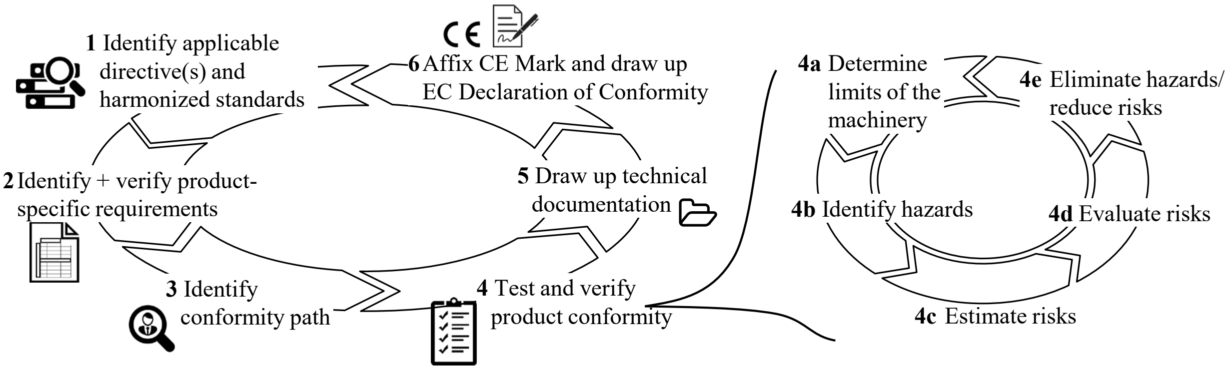
In the following sections, first the problem is presented in Section 2, followed by the current state-of-the-art approaches for risk assessment in Section 3. Section 4 presents the approach for speeding up/automating the conformity process, especially in terms of risk assessment. The approach is validated on a use case in Section 5. The results of the validation are discussed in Section 6 before an outlook on future research work in this field is given in Section 7.

1. Description of the problem

One of the questions that arises in the adaptable concept of the MFF is how to ensure regulatory safety demands. In conventional production, the manufacturer of a machine has to guarantee that their design meets the safety demands of the Machinery Directive 2006/42/EC [12], while the proprietor of the facility has to fulfil the requirements according to 2009/104/EG (minimum safety and health requirements for the use of work equipment by workers at work) [13] which demands the workplace conditions and surrounding conditions to be checked if they are still consistent with the intended use according to the machine manufacturer. If the machine manufacturer does not meet these requirements according to the Machinery Directive, they may be violating the law [14] and could lose their liability insurance coverage [15]. In general, the conformity process consists of the 6 steps listed in [16] and depicted in Figure 2.

While, in production scenarios, the machine manufacturer develops in the direction of a system proprietor [17], in the concept of the MFF the proprietor develops into a machine manufacturer and also potentially fulfils both roles. Therefore, the relevance of safety engineering will grow [18] in line with the fast adaptation of machines [19]. This means that a risk assessment must be conducted after each production system modification [20], resulting in a high effort for the proprietor because now the proprietor carries out tasks which would normally be performed during the design phase [20] by experienced health and occupational safety experts [21]. If, during the risk assessment, the modification is considered as being substantial, the machine is considered a new machine and the complete conformity process must be carried out [22].

Therefore, in order to apply the ultra-flexible production system - such as the MFF - in practice as well as reconfigurable manufacturing systems, the process of checking conformity must be supported in a more targeted and automated way [21].



1. Basically, for the products considered in this approach, the main directive will be the Machinery Directive. Besides the included Low Voltage Directive, other directives are likely to be included, e.g. the EMC Directive or possibly the RoHS Directive.
2. For the Machinery Directive, the essential health and safety requirements are included in Annex I of the Machinery Directive. Additionally, the relevant harmonized standards need to be derived.
3. It has to be determined whether a notified body has to be involved. For most machinery (except that mentioned in Annex IV), an assessment of conformity with internal checks is sufficient.
4. Consists of the risk assessment procedure and verification and tests of the derived measures e.g. according to the derived harmonized standards to prove conformity of the product.
5. All relevant documents as mentioned in Annex VII must be compiled and stored. This also includes technical, electrical and geometrical specifications.
6. Declaration of Conformity must be signed by an authorized representative of the manufacturer. Additionally, the CE Mark must be affixed to the product.

Figure 2: conformity process on the basis of [16,12]

2. State of the Art

The generic risk assessment process (RAP) is defined in ISO 12100 [23]. ISO 10218-1/2 [24] extends RAP by a list of common potential hazards encountered in human robot collaboration applications, while ISO/TS-15066 provides guidelines on the bio mechanical force and pressure limits arising from transient and quasi-static collisions [25]. In this paper, we group the various RAP techniques into three main categories:

1. Informal safety analysis techniques and conventional tool
2. Model-based and formal verification methods
3. Simulation/scenario-based methods

Informal methods are general-purpose techniques used to structure RAP. Examples of such techniques are HAZard and Operability studies (HAZOP) [26], checklists [27], Failure Mode and Effect Analysis (FMEA) [28], Failure Tree Analysis (FTA) [29], and Systems Theoretic Accident Model and Processes (STAMP) [30]. These rely solely on domain expertise and human reasoning, such as HAZOP entails brainstorming meetings by interdisciplinary teams to explore possible deviations from system design, system behavior and subsequent hazards. While the later mentioned methods can only be applied to areas which have been elaborated previously with much effort, informal methods can be applied to all type of risks and production scenarios. But the engineering judgement of the risk identification and the risk evaluation is usually done with reflection on a “mind internal” comparison to comparable situations or comparable machine structure. Therefore there is a potential of not identified risks as well as under- or overestimates of risks due to cognitive biases. [31,32]

Typical **model-based methods** derive hazards from system designs using rules and formal verification methods. [33] combines HAZOP and UML to systematically identify hazards in Human Robot Collaboration (HRC) applications. [34] introduces a rule-based system to automatically link certain Product Process Resource (PPR) properties to hazards and make recommendations for corresponding safety measures. SAFER-HRC is a formal verification method that guarantees the safety of the given HRC application [35].

[36] introduces a meta-model to automate a design space exploration which allows measuring impacts on design changes automatically. To increase automation in risk assessment [37] presents a model-based and software-assisted approach, performing hazard analysis at runtime on single components. While model-based methods can automatically identify hazards and formally verify the safety of an HRC application, they require rigorous formal models of the application – a highly work-intensive and knowledge-intensive task - and are only as complete or correct as the underlying model and verification rules.

Few **simulation-based methods** exist, inter alia, for identifying hazards [38] and the optimized placement of Rapid Response Mechanisms (RRM) [39] in HRC systems, for Automated Guided Vehicles (AGVs) in automated warehouses [40], or for cobots in domestic environments [41]. [40] combines a simulation-based approach with HAZOP, however with a focus on automated warehouses with AGVs. The authors simulate advanced sensors functionalities, such as depth cameras, which cannot be easily formulated by model-based and formal methods. Huck et al. [38] address hazard identification using optimization-based methods which search for risky human behavior that could lead to hazards. In a proof of concept, Huck et al. are able to identify hazards arising from unplanned transient contacts. However, the approach is not guaranteed to converge in a polynomial time. Furthermore, due to simplification of the human model in order to utilize a *branch&bound* algorithm, this approach can only be used to falsify specific safety configurations which are related to specific hazards. It is unable to guarantee safety because hazards may be overlooked.

Several contributions published in the production domain address an automation of single process steps in the CE conformity process for reconfigurable systems. Some work has been done on automating the identification of hazards associated with specific machine modules. Identification of hazards and evaluation of safety at system level still relies on expert knowledge "[originating] from a human integrator" [42]. [36,37,42]

The impact of modifying the configuration of modular machines - consisting of several single components – on the validity of safety configurations and subsequently the CE mark has not been addressed in detail. Specifically, the interactions between safety components and functional components being added or removed have not been in the focus of those publications.

In this publication the challenges of declaring CE conformity of modular machines are discussed. Following that, an approach for accelerating the declaration process is presented, considering current directives and standards. Consequently, all steps of the CE conformity process are considered - from single component up to system level - and their feasibility as well as degree of automation are discussed.

3. CE Configuration

As specified in the description of the problem, the constant demand-driven modification of machinery results in significant work for the proprietor in order to guarantee compliance with the machine's standards and directives. During the conformity assessment procedure, the machine may not be productive, leading to downtime and lost added value (2006/42/EC). The presented approach includes a framework to help the proprietor with the conformity process and automate the respective steps. As a result, time-

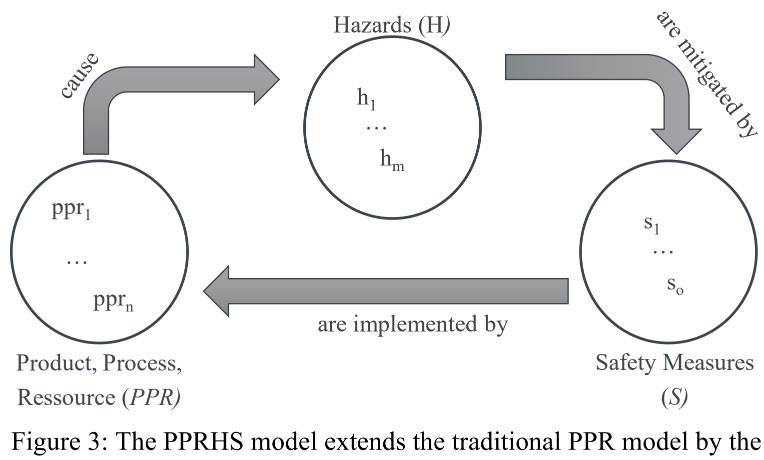


Figure 3: The PPRHS model extends the traditional PPR model by the hazards caused by the product, process or resource features of a workplace, as well as the mitigating safety measures that in turn are also implemented by PPR features.

consuming conformity assessment procedures can be avoided and downtime and costs for modular machinery minimized. The framework's contribution to Step 4 of the conformity process, which is depicted in Figure 2, is based on three core elements: 1) database of CE compliant configurations of machinery; 2) selection algorithm for identifying suitable machine configurations that meet process/product requirements, and 3) a reasoning algorithm that uses the PPRHS (**P**roduct, **P**rocess, **R**esource, **H**azard, **S**afety measure) model depicted in Figure 3, to determine whether a new configuration contains any unmitigated risks.

The used **database** contains a) the available machine modules for building modular machines, b) the CE compliant configurations of the used modules and their data. Furthermore, the saved data for each CE compliant configuration includes:

- machine-readable documentation of the configuration of machinery [43–45] (e. g. links between modules, geometrical, electrical wiring plans)
    - required capabilities of the configuration
    - available capabilities of the configuration
- standards used at the time of conducting CE compliancy (e. g. DIN EN ISO 10218-2:2012-06)
- family tree and timeline of configurations and sub-configurations used to modify the configuration
- risk assessment, including identified hazards and selected safety measures
- declaration of conformity

If a configuration is declared as being conform, the above-mentioned data is saved, thus extending the database. In general, modifications should only be implemented if they benefit value-adding (e.g. integration of necessary new capabilities or removal of obsolete ones).

To derive the required manufacturing capabilities from product and process requirements and to select configurations in the database which fulfil these requirements, a **selector algorithm** is used. The selector searches the database and compares the necessary derived capabilities with the available capabilities of the configurations in the database and outputs either a CE-compliant configurations that fulfils all requirements or a "best match". That best match is then further modified by the engineer in terms of **P**roduct, **P**rocess and **R**esource features to fulfils all requirements.

When a configuration is modified, risks and adequacy of safety measures must be verified, as:

- Existing hazards can become unmitigated (risk level increases) due to a modification/elimination of the PPR features implementing the corresponding safety measures.
- New but known hazards, i.e. safety measure is also known, may manifest due to a modified PPR features
- New unknown hazards, i.e. associated safety measure is not yet associated, may manifest due to new PPR features

A **reasoning algorithm** following formal rules and utilizing a simulation-based risk assessment procedure determines if the hazard increases and safety measures are sufficient.

This paper summarizes the framework's outcome for two different conformity process scenarios, outlined in Figure 4.

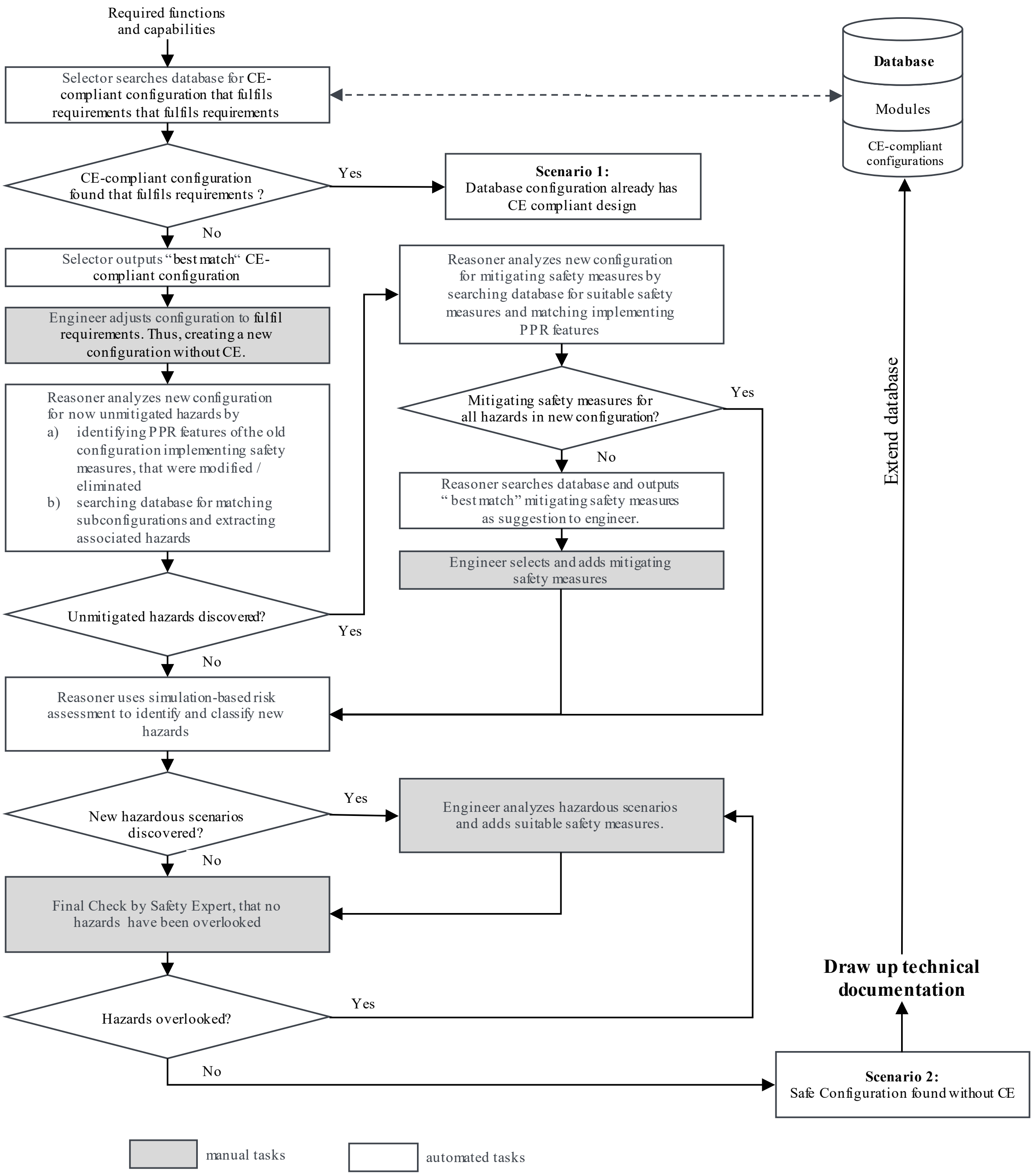


Figure 4: Outcome for different conformity process scenarios

Scenario 1: The selector finds one or more CE-compliant configurations in the database which fulfil all the requirements.

- These configurations can be reused as the "new" configuration. Since the new configuration, the configuration in the database, and the used standards are identical, the documentation of the database configuration can be reused for Step 5 of the conformity process.

Scenario 2: The configuration found is a modified configuration of a previous CE-compliant configuration (stored in database). In that case, a risk assessment procedure consisting of identifying hazards and selecting suitable safety measures is mandatory. To identify hazards the following four evaluation steps are conducted:

- In a first step, the reasoner analyzes whether hazards of the previous CE-compliant configuration become unmitigated due to elimination or modification of PPR features implementing safety measures. It should be noted that the elimination or modification of PPR features could also lead to the elimination or risk decrease of hazards.
- In a second step, the reasoner searches the database for matching sub-configurations and the reasoning algorithm uses the PPRHS model to evaluate adequacy of the safety measures and no increase in risk.
- In a third step, simulation-aided risk assessment (SARA) tool is used to identify potential new hazards that result from the "sum of all sub-configurations". The tool uses a digital twin of the machine and an adversarial human digital model to simulate different interaction behaviors in a physics simulation, where the human model is "trying to get hurt" and the physics engine is used to evaluate resulting collision forces and pressures.
- Since the rules and SARA tool used by the reasoner cannot claim completeness – they can only falsify safety; not verify it - a final check (step 4) is performed by a safety expert to make sure that no hazards are overlooked.

In case an unmitigated or new hazards (see Figure 5) in any of the previous four steps is identified, three distinct cases can be distinguished:

- Case I: The PPR features causing those hazards can be identified as a sub-configuration in the database (common for results from step 1 and 2 of the hazard identification process), meaning that those hazards have been encountered before and suitable safety measures exist. It is then just a matter of identifying the PPR features implementing those safety measures in the configuration or adding them to the configuration if missing.
- Case II: The PPR features causing those hazards are similar but not an exact match to a sub-configuration in the database. The reasoner can output "best match" safety measures, but it is up to the engineer to evaluate, whether safety measure option is a) adequate or b) also requires modification.
- Case III: New unknown hazards (common for results from step 3 and 4 of the hazard identification process). In that case, the engineer needs to codify the underlying PPR features and suitable safety measures.

In Case I, the risk assessment procedure can be quickly passed. Further, the safety documentation which is already available can be largely reused in Step 5 of the conformity process. Case IIa requires more effort, but again a lot of the existing documentation can be reused, or the lacking documentation about the new mapping of hazards to safety measures can be simply added. In cases IIb and III, new previously unknown hazards are identified and no adequate safety measures exist. Thus, requiring a lot of the effort engineer, causing machine downtime and added human effort.

Therefore, if those cases can be avoided, which is one of the framework's core objectives, ultra-flexible production systems will be more efficient. To guarantee sufficient human safety, the risk must be reduced until it is within the limits defined by the standards. The documentation of this risk assessment is then added to the documentation of the machine. Furthermore, the new CE-compliant configuration extends the framework's database.

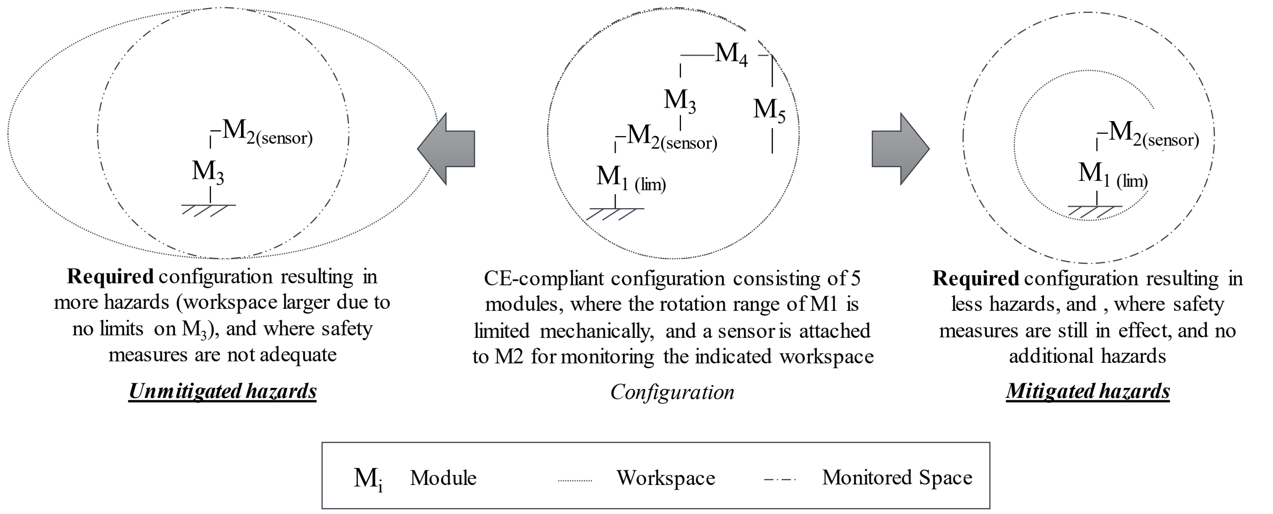


Figure 5: Examples of modifications of CE-compliant configuration (middle) resulting in reduced hazards / risk levels (right) or in unmitigated hazards / risks (left)

The configured machine still has to be approved (conducted manually) in order to avoid modification errors, such as wrongly-attached wires, sensors or the omission of electrical grounding. The construction must correspond to the CE-compliant design in the database. If more than one configuration is available after the database search, the found configurations are prioritized according to multi-criterial objectives, such as time, cost, or associated risks. Reusing configurations in the database enables Steps 4 to 6 in the conformity process to be accomplished quickly, thus minimizing machine downtime.

It should also be noted that the CE-conform configurations in the database can only be employed if the relevant standards are still valid. Otherwise, their CE marking can be used by the reasoner as a valid option before it is updated according to current standards during reconfiguration.

4. Use Case

Using the modular, mobile machines (see Figure 6) designed by IFF, the presented framework are applied to an example of a scenario. In this scenario a CPS, represented by a mobile modular machine consisting of a modular robot, an automated guided vehicle (AGV), and its respective power supply, is assigned an assembly task. This CPS was digitally modeled and visualized in Unity Engine and is continuously updated regarding the current configuration. This game engine allows to integrate 3D models, scripting and several relevant technologies, like communication network protocols. Using Unity Engine and physics engines, like MuJoCo, it is possible to simulate CPS and their physical properties, including collisions and mechanical behavior according to the current configuration [46–48]. The assembled product is expected to be delivered to a specific location in the factory. The transport time is used to assemble the product. At this stage, the robot is in a 3-axis configuration. However, the production engineer discovers that the robot is unable to fulfil the planned assembly task in its current configuration.

After an iterative process, the worker derives that the minimal required configuration to achieve the task needs the integration of three additional modules. He or she modifies the robot accordingly. Following the procedures detailed in Section 3, the machine's current configuration must comply with ISO 10218-2 [24]. IFF is considered as the machine's integrator and is accordingly responsible for compliance with the Machinery Directive as well as with the specific robot standard [24]. On conducting a database search using the framework presented, neither a configuration nor sub-configuration fulfilling the requirements can be found. It cannot be proven that the new configuration poses a lower risk than the old. Additional safety measures are required and a "classical" CE process is performed. The resulting documentation and the declaration of CE conformity is then stored depending on the time and location, and the production engineer is then permitted to affix the CE marking. Following the procedure described, the modification process is complete and the CE-conform modular machine is now able to start its assembly task. Further, the new configuration is stored in the described database. When the scheduled task is finished, the worker removes one of the modules again because it is needed for another robot. Utilizing the framework presented in Section 3, it can be determined that the configuration is a subclass of a CE compliant configuration with the same safety measures and no additional hazards identified (see Figure 5). Consequently, it can be proven that the new configuration poses no higher risk than the preceding configuration. Thus, the safety measures are considered to be adequate. The configuration only has to be approved and the technical documentation added.

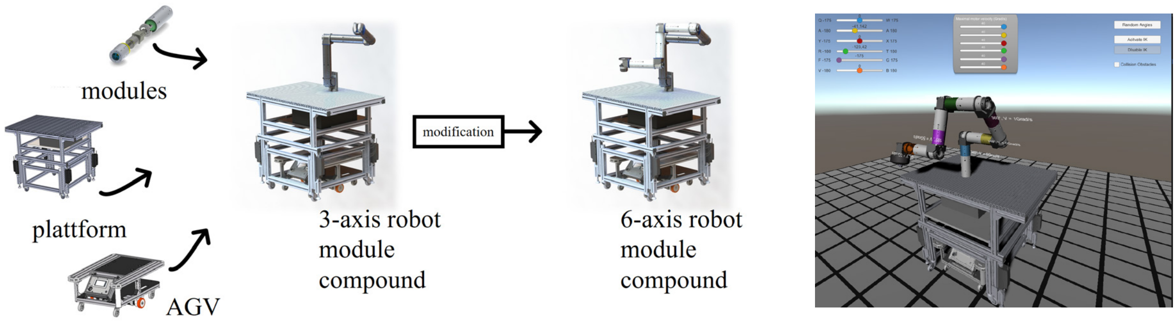


Figure 6: Modification of a machine consisting of different modules (left) and its digital representation (right) in Unity Engine

Later, the worker needs the robot to accomplish a different task. Using the framework once again, an already CE compliant design suitable for the task can be found. The configuration therefore only has to be approved and the appropriate technical documentation compiled.

5. Discussion of results

The presented framework proposes an approach to minimize downtime by speeding up and automating the respective conformity assessment procedures.

The analysis in this paper is limited to machines falling under the machinery directive (2006/42/EC), excluding the machinery listed in Annex 4 of the directive. Due to this limitation, the only applicable directive identified is the machinery directive excluding Annex 4, i.e. excluding machines requiring conformity assessment by a notified body. Accordingly, this publication only considers machines where the conformity process requires a self-declaration.

Steps 1, 2, 3 and 5 in the CE marking process can already be automated with database software which output the respective applicable harmonized standards, including the requirements of the machinery directive. Step 4 includes the database search and case distinction into three scenarios.

Scenario 1 can be fully automated. The degree of process automation in Scenario 2 depends on the completeness of the reasoning algorithm’s rules for determining that there is no increased risk and that safety measures are sufficient. Gaps identified in the rule system can be filled by adding more rules (which would result in the common problem of keeping rule-based systems up to date) or by integrating additional tools, e.g. simulation models, which can be used to falsify safety measures. While this concept will never yield a 100 percent guarantee, it could be argued that the results obtained with it will approach and surpass the results of the human safety expert.

Scenario 2 and 3 require a conformity assessment procedure which not only involves human effort for Steps 4 to 6 but also causes machine downtime. However, as the framework matures and is filled with more CE compliant configurations, Scenario 3 Case IIb and Case III in scenario 2 become less and less likely. Applied to production systems and "self-extending" networks around the globe, cases like Scenario 3 could become insignificant and the conformity process almost fully automatable. However, going forward, one challenge will be the necessary processing power, depending on the number of modules and configuration possibilities (framework scale up).

As the frequency of system modification increases, the proportion of modification-time = downtime = lost added value in today's machine lifecycle will increase, resulting in decreasing OEE. At best, the framework aims at eliminating the time required for the conformity process, which is considered part of modification time.

Related to machines that are modified within minutes to hours and extrapolated to hundreds of machines in a production system, the use of this framework could save considerable modification time that could be used for manufacturing. In addition, the framework evaluates whether a conformity assessment procedure is required and offers the additional criterion of whether the modification will add value.

6. Outlook

Ongoing activities at IPA and IFF are filling the gaps in the rule-based system by restructuring the guidelines, as well as further **integrating a simulation-based risk assessment module**, which uses a digital twin of the machine configuration and erroneous human behavior to attempt to falsify safety measures. Furthermore, to **quantify the achieved time savings**, the framework will be benchmarked using standardized use cases that reflect the typical reconfiguration scenario of the future factory. Hence, the definition of such benchmark use cases is a challenge within itself.

Another research avenue being pursued is connecting the framework to the digital safety shadow of machines. Like the digital process shadow, the **digital safety shadow** primarily gathers safety-related data over time and contextualizes these data into usable information. Such information could be used to quantify risk frequency and probabilities in order to either verify or reassess the risks associated with an HRC application. For example, the information on how often a worker enters a designated hazard area could be used to confirm the risk frequency of the hazard in the risk assessment. If the risk frequency is much higher, additional safety measures will be required; if it is much lower, safety measures can be relaxed and the flexibility of the application increased.

Another point to be examined in the future is the fact that currently no additional specific formal qualification of the CE declarant is required for the declaration of CE conformity. This raises the question of how a person without specific training can assess hazards. Taking into account typical definitions of competence in production technology [49–52], the analysis of the situation can only be done by including and estimating hazards of already known systems. An estimation via similarities of known systems and the abstraction to common features represent a form of modelling. In this context, the question is whether and if so, how human abstraction and modelling can be transferred to a decision support approach to further accelerate and improve decision making.

References

[1] Koren, Y., Shpitalni, M., Gu, P., Hu, S.J., 2015. Product Design for Mass-Individualization. Procedia CIRP 36, 64–71.

[2] Chryssolouris, G., 2013. Manufacturing systems: theory and practice. Springer Science & Business Media.

[3] Mehrabi, M.G., Ulsoy, A.G., Koren, Y., 2000. Reconfigurable manufacturing systems: Key to future manufacturing. Journal of Intelligent Manufacturing 11 (4), 403–419.

[4] Wiendahl, H.-P., ElMaraghy, 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.

[5] Cardin, O., 2019. Classification of cyber-physical production systems applications: Proposition of an analysis framework. Computers in Industry 104, 11–21.

[6] 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.

[7] Stock, D., Bauernhansl, T., Weyrich, M., Feurer, M., Wutzke, R., 2020 - 2020. System Architectures for Cyber-Physical Production Systems enabling Self-X and Autonomy, in: 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria. 08.09.2020 - 11.09.2020. IEEE, pp. 148–155.

[8] Schlegel, T., Siegert, J., Bauernhansl, T., 2019. Metrological Production Control for Ultra-flexible Factories. Procedia CIRP 81, 1313–1318.

[9] Siegert, J., Schlegel, T., Groß, E., Bauernhansl, T., 2017. Standardized Coordinate System for Factory and Production Planning. Procedia Manufacturing 9, 127–134.

[10] Siegert, J., Schlegel, T., Bauernhansl, T., 2018. Matrix Fusion Factory. Procedia Manufacturing 23, 177–182.

[11] Liggesmeyer, P., Trapp, M., 2014. Safety: Herausforderungen und Lösungsansätze, in: Bauernhansl, T., Hompel, M. ten, Vogel-Heuser, B. (Eds.), Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung \textperiodcentered Migration. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 433–450.

[12] Office, P., 2009. Directive 2009/104/EC of the European Parliament and of the Council of 16 September 2009 concerning the minimum safety and health requirements for the use of work equipment by workers at work (second individual Directive within the meaning of Article 16(1) of Directive 89/391/EEC).

[13] Neudörfer, A., 2016. Konstruieren sicherheitsgerechter Produkte: Methoden und systematische Lösungssammlungen zur EG-Maschinenrichtlinie, 7., aktualisierte Auflage 2016 ed. Springer Vieweg, Berlin, Heidelberg, 669 pp.

[14] Wollert, J., 2015 - 2015. Wireless systems for machinery safety, in: 2015 16th International Conference on Research and Education in Mechatronics (REM). 2015 16th International Conference on Research and Education in Mechatronics (REM), Bochum, Germany. 18.11.2015 - 20.11.2015. IEEE, pp. 88–91.

[15] Jespen, T., 2016. Risk Assessments and Safe Machinery. Springer International Publishing, Cham, 263 pp.

[16] MEMO/10/257 European Union, 2010. CE marking – what does it really mean?

[17] Westkämper, H., 1997. Manufacturing on demand in production networks: CIRP Annals-Manufacturing Technology (46(1),), 329–334.

[18] Markku Mattila, Maarit Perälä and Vesa Vannas, 1996. Flexible Manufacturing Systems' compliance to the safety standards: International Journal of Advanced Manufacturing Technology.

[19] Markis, A., Papa, M., Kaselautzke, D., Rathmair, M., Sattinger, V., Brandstötter, M., 2019. Safety of Mobile Robot Systems in Industrial Applications, 7 pp.

[20] Ina Schaefer, Loek Cleophas, Michael Felderer Ina, 2018. Challenges in the engineering of adaptable and flexible industrial factories: Modellier Lecture Notes in Informatics (LNI) Gesellschaft für Informatik Bonn 2018 ung in der Entwicklung von kollaborativen eingebetteten Systemen (MEKES).

[21] Koo, C.H., Vorderer, M., Junker, S., Schröck, S., Verl, A., 2018. Challenges and requirements for the safety compliant operation of reconfigurable manufacturing systems. Procedia CIRP 72, 1100–1105.

[22] ANASTASI, S., MONICA, L., 2017. EVOLUTION OF EUROPEAN PRODUCT DIRECTIVES IN PERSPECTIVE OF INDUSTRY 4.0, in: Safety and Security Engineering VII. SAFE 2017, Rome, Italy. 06.09.2017 - 08.09.2017. WIT PressSouthampton UK, pp. 163–168.

[23] ISO 12100:2010. Safety of machinery — General principles for design — Risk assessment and risk reduction.

[24] DIN e.V. DIN EN ISO 10218-2 Industrieroboter - Sicherheitsanforderungen - Teil 2: Robotersysteme und Integration. Beuth Verlag, 93 pp.

[25] ISO/TS 15066:2016. Robots and robotic devices — Collaborative robots.

[26] Siddiqui, D.N., Nandan, A., Sharma, M., Srivastava, A., 2014. Risk management techniques HAZOP and HAZID study. Int J Occup Health Saf, Fire Environ Allied Sci 1 (1), 5–8.

[27] Hunt, E.R., Hauert, S., 2020. A checklist for safe robot swarms. Nat Mach Intell 2 (8), 420–422.

[28] IEC 60812:2018. Failure modes and effects analysis (FMEA and FMECA). VDE.

[29] IEC 61025:2006. Fault tree analysis (FTA).

[30] Lower, M., Magott, J., Skorupski, J., 2018. A System-Theoretic Accident Model and Process with Human Factors Analysis and Classification System taxonomy. Safety Science 110, 393–410.

[31] Baybutt, P., 2018. The validity of engineering judgment and expert opinion in hazard and risk analysis: The influence of cognitive biases. Proc. Safety Prog. 37 (2), 205–210.

[32] Jong, H. de, Blom, H., Stroeve, S., 2007. How to identify unimaginable hazards?

[33] Guiochet, J., 2016. Hazard analysis of human–robot interactions with HAZOP–UML. Safety Science 84, 225–237.

[34] Awad, R., Fechter, M., van Heerden, J., 2017 - 2017. Integrated risk assessment and safety consideration during design of HRC workplaces, in: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol. 12.09.2017 - 15.09.2017. IEEE, pp. 1–10.

[35] Askarpour, M., Mandrioli, D., Rossi, M., Vicentini, F., 2016. SAFER-HRC: Safety Analysis Through Formal vERification in Human-Robot Collaboration, in: Skavhaug, A., Guiochet, J., Bitsch, F. (Eds.), Computer Safety, Reliability, and Security, vol. 9922. Springer International Publishing, Cham, pp. 283–295.

[36] Hofig, K., Klein, C., Rothbauer, S., Zeller, M., Vorderer, M., Koo, C.H., 92019. A Meta-model for Process Failure Mode and Effects Analysis (PFMEA), in: 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). 2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain. 10.09.2019 - 13.09.2019. IEEE, pp. 1199–1202.

[37] Koo, C.H., Schröck, S., Vorderer, M., Richter, J., Verl, A., 2020. A model-based and software-assisted safety assessment concept for reconfigurable PnP-systems. Procedia CIRP 93, 359–364.

[38] Huck, T.P., Ledermann, C., Kroger, T., 2020 - 2020. Simulation-based Testing for Early Safety-Validation of Robot Systems, in: 2020 IEEE Symposium on Product Compliance Engineering - (SPCE Portland). 2020 IEEE Symposium on Product Compliance Engineering - (SPCE Portland), Portland, OR, USA. 16.11.2020 - 17.11.2020. IEEE, pp. 1–6.

[39] Bobka, P., Germann, T., Heyn, J.K., Gerbers, R., Dietrich, F., Dröder, K., 2016. Simulation Platform to Investigate Safe Operation of Human-Robot Collaboration Systems. Procedia CIRP 44, 187–192.

[40] Inam, R., Raizer, K., Hata, A., Souza, R., Forsman, E., Cao, E., Wang, S., 2018 - 2018. Risk Assessment for Human-Robot Collaboration in an automated warehouse scenario, in: 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA). 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), Turin. 04.09.2018 - 07.09.2018. IEEE, pp. 743–751.

[41] Kaonain, T.E., Abdul Rahman, M.A., Mohammed Ariff, M.H., Mondal, K., 2020. A simulated risk assessment of human-robot interaction in the domestic environment. IJRA 9 (4), 300.

[42] Moncada, D.S.V., Schneider, D., Petrovska, A., Laxman, N., Möhrle, F., Rothbauer, S., Zeller, M., Koo, C.H., Safdari, S., 2021. Dynamic Safety Certification for Collaborative Embedded Systems at Runtime, in: Böhm, W., Broy, M., Klein, C., Pohl, K., Rumpe, B., Schröck, S. (Eds.), Model-Based Engineering of Collaborative Embedded Systems. Springer International Publishing, Cham, pp. 171–196.

[43] Fechter, M., Keller, R., Chen, S., Seeber, C., 2019. Heuristic Search Based Design of Hybrid, Collaborative Assembly Systems, in: Advances in Production Research, Cham. 2019. Springer International Publishing, Cham, pp. 188–197.

[44] Fechter, M., Seeber, C., Chen, S., 2018. Integrated Process Planning and Resource Allocation for Collaborative Robot Workplace Design. Procedia CIRP 72, 39–44.

[45] Zander, S., Awad, R., 92015. Expressing and reasoning on features of robot-centric workplaces using ontological semantics, in: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany. 28.09.2015 - 02.10.2015. IEEE, pp. 2889–2896.

[46] Siegert, J., Zarco, L., Schlegel, T., 2020 - 2020. Universal Accessibility Concept for Controlling Production Means in Manufacturing Systems, in: 2020 IEEE International Conference on Industrial Technology (ICIT). 2020 IEEE International Conference on Industrial Technology (ICIT), Buenos Aires, Argentina. 26.02.2020 - 28.02.2020. IEEE, pp. 349–354.

[47] Siegert, J., Zarco, L., Schlegel, T., Bauernhansl, T., 2020. Software Control System Requirements for Ultra-Flexible Learning Factories. Procedia Manufacturing 45, 442–447.

[48] Zarco, L., Siegert, J., Schlegel, T., Bauernhansl, T., 2020. Determining and Evaluating Trajectories of a Modular Machine End-Effector in a Manufacturing Environment using a Game Engine. Procedia CIRP 93, 1079–1084.

[49] Erpenbeck, J., Rosenstiel, L., 2017. Handbuch Kompetenzmessung, 3rd ed. Schäfer Poeschel, 12 pp.

[50] North, K., Reinhardt, K., Sieber-Suter, B., 2013. Kompetenzmanagement in der Praxis: Mitarbeiterkompetenzen systematisch identifizieren, nutzen und entwickeln ; mit vielen Fallbeispielen, 2., überarb. u. erw. Aufl. 2013 ed. Gabler Verlag, Wiesbaden, 293 pp.

[51] Siegert, J., Schlegel, T., Bauernhansl, T., 2020. Verifiable Competencies for Production Technology. Procedia Manufacturing 45, 466–472.

[52] Thomas Jackwerth, Angela Jäger, Sandra Güth - Fraunhofer ISI, 2018. Fähigkeiten und Kompetenzen gezielt ausbauen - Strategische Kompetenzentwicklung für die Produktion. Modernisierung der Produktion (73).

Biography

Jörg Siegert (*1978) has been the Senior Engineer and Deputy Head of the Institute of Industrial Manufacturing and Management IFF since 2012: He has also been acting deputy head of the newly founded Institute of Energy Efficiency in Production EEP since 2012.

Lukas Krispin (*1995) studied mechanical engineering at the University of Stuttgart and has been working as a research associate at the Institute of Industrial Manufacturing and Management (IFF) since 2020.

Awad Ramez (*1984) obtained his diploma degree with honors in industrial engineering from the University of Karlsruhe in 2008. He worked from 2008 - 2012 as a researcher at the Karlsruhe Institute for Technology (KIT) in the area of Digital Factory. In 2012 he joined the Department for Robot and Assistive Systems at Fraunhofer IPA, where he heads the research group for Assembly Automation since 2017.

Mohamed El-Shamouty (*1994) received his B.Sc. degree with honors in Information Engineering and Technology from the German University in Cairo and his M.Sc. degree in Information Technology from the University of Stuttgart. His main focus is making robots cognitive and safe with the aid of machine learning for industrial applications.

Thilo Schlegel (*1985) graduated in industrial engineering at the University of Stuttgart and is working as a research assistant for Fraunhofer Institute for Manufacturing Engineering and Automation IPA and the Institute of Industrial Manufacturing and Management IFF where he heads the Matrix Fusion Factory since 2019.

Florian Roth (*1995) studied mechanical engineering at the University of Stuttgart and has been working as a research associate at the Institute of Industrial Manufacturing and Management (IFF) since 2020.

Oliver Mannuß (*1969) studied mechanical engineering at the University of Stuttgart and heads the research group of Quality- and Reliability Management since 2006 and held the additional role as the CE-Coordinator of the Fraunhofer IPA. His main research focus are risk management methodologies during product and production process development.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Digital Industrial B2B Platform Patterns From A Business Perspective

Fabian Hartner[1], Ulrich Löwen[2], Jörg Franke[1]
[1]Institute for Factory Automation and Production Systems, University of Erlangen-Nürnberg, Germany
[2]Siemens AG Technology, Erlangen, Germany

Abstract

Platform companies such as Amazon and Google have changed the business-to-consumer (B2C) sector at a rapid pace. In the context of the digitization of products, production systems, and processes, the importance of digital business-to-business (B2B) platforms is also increasing. In the manufacturing industry, value chains change to complex platform-based value networks between business entities. This publication presents the results from the analysis of more than 160 platforms in the manufacturing industry from a business perspective. Using a value network-oriented methodology, defining business roles as well as revenue streams and business relationships, recurring value-creating interactions were observed. In total, eight clearly distinguishable platform patterns have been identified. In this article, these patterns are described in detail with their value and revenue streams as well as their benefits for the individual roles. These patterns aim to achieve a well-founded understanding of platforms in the manufacturing industry and form the fundamentals for further analysis for successful digital industrial B2B platforms.

Keywords

Digital Platforms; Value Networks; Platform Patterns; Business Models; Manufacturing Industry;

1. Introduction

In the ongoing digitalisation, in addition to new technical developments, new digital business models are emerging. Since very successful platform companies have been developed in the business-to-consumer (B2C) market (e.g. Apple, Uber), platform-based business models have attracted a high level of attention from individuals, companies, politics and researchers. Following the successes in B2C, digital platforms and associated business models are also becoming increasingly relevant in the business-to-business (B2B) environment and especially in the manufacturing industry, where more and more platform-based offerings are coming onto the market. According to Kenney et al. [1], digital platforms represent powerful principles for economic effects and designs in the manufacturing industry in the coming decades. Pauli et al. shows in [2] that industrial platforms are different from other platforms already intensively considered in previous studies (e.g. mobile platforms, video games consoles). Therefore, previous results from other domains cannot be directly transferred without further reflection. In the manufacturing industry, an increasing number and a wide variety of platforms can be observed that offer users a diverse range of value propositions. In order to understand and explain the underlying dynamics of the very heterogeneous platforms, digital platforms can be classified into different types [3]. This publication shows the results of the analysis of practical examples using a methodology concentrating on value and revenue streams from a business perspective (see chapter 3) and describes identified pattern of digital B2B platforms in the manufacturing industry (see chapter 4).

DOI: https://doi.org/10.15488/11247

2. Theoretical Background and related work

This publication focuses on platforms in the manufacturing industry. According to the United Nations ISIC classification system [4], manufacturing refers to an industry in which a physical or chemical transformation of materials, substances or components into new products takes place typically in plants or factories with the use of powered machinery and material handling equipment. Furthermore, this research focuses on platforms facilitating B2B transactions. According to Kleinaltenkamp and Saab [5], the B2B market refers to a business transaction between two or more companies where the customer typically further processes the purchased product or uses the product for its own service or production process. Transactions typically are characterised by the involvement of several people and departments on supplier and customer side and by the often complexity of the services. According to Falck and Koenen [6], the differences between B2B and B2C are particularly apparent with regard to the size of transactions, customer relationships, focus on profits and efficiency, as well as contract terms and lead to different mechanisms for B2B platforms compared to B2C.

The term platform cannot be defined uniformly. In order to define the term platform, a distinction should be made between two different perspectives on platforms. On the one hand, Gawer [7] describes a technical understanding in which platforms are technological architectures that support companies with modules in the development and innovation of modular products or technologies. On the other hand, an economic perspective defines a platform according to Parker et al. [8] as "a business based on enabling value-creating interactions between external producers and consumers". This perspective on business interactions between external partners is also adopted in this publication by focussing on value-creating relationships between business partners based on digital technologies.

Regarding different platform types, some general classifications have already been developed. Engelhardt et al. [9] divide platforms very generic into transaction-centred platforms, which act as intermediaries to bring offer and demand together, and data-centred platforms, which connect data based on hardware and software. Täuscher and Laudien [10] focus on marketplaces as business models and identified six different marketplace platform types across industries. Evans and Gawer [3] distinguish between four platform types: transaction platforms that facilitate transactions between different users, innovation platforms as a foundation on which companies develop complementary offerings, integrated platforms as a combination of transaction and innovation platforms and investment platforms that consist of several companies acting as investors. Due to the complexity of industrial platforms, a more detailed level of abstraction is necessary.

However, these three studies follow very general distinctions and pay insufficient attention to B2B and industry-specific platform offerings. An initial analysis with detailed focus on industrial platforms has so far been carried out by Wortmann et al. [11], in which the following five platform types were defined: Two- and multi-sided markets as intermediaries for matching between actors, service platforms as intermediaries offering targeted collaboration, IoT platforms offering technical solutions in pipeline models, IoT-based intermediaries use IoT platforms for intermediary functions and Smart IoT platforms as an extension of IoT platforms with further services. At the current state of the art, there is no detailed research that defined their classification based on value propositions, value creation relationships and income mechanisms in industrial B2B. For this reason, this publication addresses the research question, which B2B platform patterns can be distinguished in the manufacturing industry based on the value creation relationships.

3. Methodology and research procedure

To ensure a systematically consistent approach from a business perspective, this research applied a value creation perspective to examples of digital platforms in the manufacturing industry. A value network perspective with the analysis of physical and service value streams was chosen to analyse the business transactions including benefits, value propositions, and revenue flow between partners through offerings from a platform provider. The methodology was published by the working group "Digital Business Models in Industry 4.0" of Plattform Industrie 4.0 in [12] and initially applied to five examples. Until yet, the methodology has been proven and established in several publications for the analysis of practical examples on a uniform business level, as in [13] and [14]. The representation of the value creation networks follows a methodology in which nodes (as icons) represent certain business roles of a company and lines (as arrows) represent corresponding value creation relationships. The value creation relationships are divided into physical value creation and value creation through services, and the concrete proposition is supplemented by a brief caption. The representation of the revenue streams is done with dashed and labelled arrows and shows who pays for which service or product.

For a comprehensive research basis, the methodology was applied to platform examples from the manufacturing industry. In this work, platforms are considered that have at least one actor in their value networks that acts according to the definition of ISIC and has an industrial character (e.g. through machine use, work in factories). The following research procedure was used to select platform examples as a data basis for further research and analysis. First, publications focusing on platforms in the manufacturing industry were searched for concrete business examples (e.g. [13], [15]). This was expanded by means of an internet search via Google and Google Scholar with the terms "digital B2B platform", "platform businesses manufacturing industry" and related terms. An example was included in the research base if it was described with the term "platform" by the provider or the author, considers primarily B2B transactions, was available on the market (no research projects) and was related to the manufacturing industry. Based on these criteria, 160 examples were identified in a first research, which have no claim to completeness. In a second step, business roles as well as their business relationships and their value propositions were defined and visualized with the value network methodology. Furthermore, the revenue streams between the business roles were analysed and documented. The information in this step was primarily taken from published company descriptions and partly expanded by contacting and interviewing the platform operators.

After a clear understanding of the platform examples and their value networks, the individual examples were analysed in a third step for recurring value relationships and service offerings based on their value networks and revenue streams. Generalised, eight clearly separable platform patterns from a business perspective could be identified, which are described in detail in the following chapter 4. A draft of the first three patterns has already been presented in [16].

4. Industrial B2B platform patterns

In the following chapter, the eight identified digital B2B platform patterns are presented and described with their respective value streams and revenue streams. In addition, a description is given of what the platform provider specifically offers and how each role benefits from it. Each pattern is based on specific practical examples. The patterns represent abstracted value creation networks and revenue streams, each of which can be modified in company-specific variants. The value streams focus on the platform operator and the roles receiving value propositions directly from the platform operator.

4.1 Innovation platform pattern

The innovation platform operator (shown in Figure 1) has the business purpose of enabling connectivity between machine users for enabling complementary services and solutions. To enable and support the analysis of usage information, service provider can use services of the innovation platform to develop applications (apps). These apps can be executed on the platform and have access to provided information by the machine user (depending on access rights). Practical examples are Siemens MindSphere, Bosch IoT Suite or GE Predix. The innovation platform pattern does not focus on mediating apps of 3rd parties via an app store.

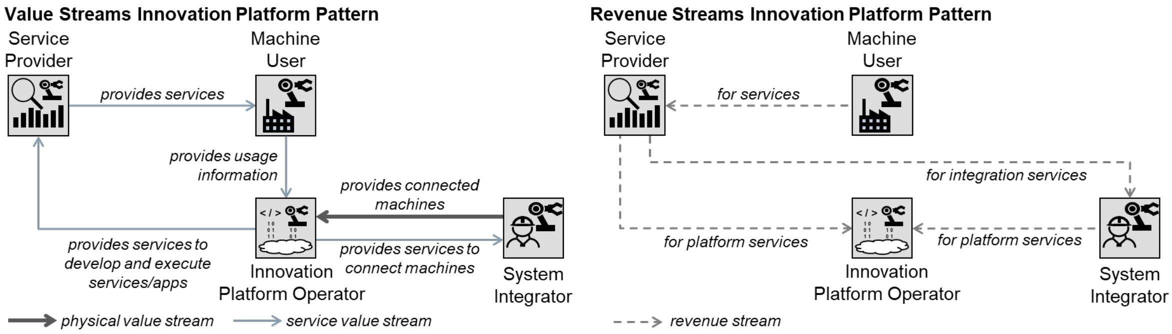


Figure 1: Value and revenue streams of the innovation platform pattern

The innovation platform operator provides a technical infrastructure that enables users to obtain usage information from machines and receive process-optimising services. The innovation platform operator enables other roles to offer and monetise new services. The business partners benefit through following aspects:

- The service provider benefits by being able to offer new or previously not economically possible services to the machine operator through the available usage information of the machines and services offered by the innovation platform operator.
- The machine user benefits by optimising internal processes by receiving services (e.g. machine downtime prediction, quality monitoring).
- The system integrator enables machines to transmit usage information and adapt the innovation platform for the specific requirements of the machine user by using services of the innovation platform operator.

The business roles involved in the value network receive revenue for the following value propositions:

- The innovation platform operator receives revenue from the service provider who pay according to the platform usage (e.g. number of connected machines) and from the system integrator for the platform services.
- The service provider receives revenues according to contractual agreements for provided services.
- The system integrator receives revenue for its integration services to connect machines.

4.2 Brokerage platform pattern

The purpose of the brokerage platform operator (shown in Figure 2) is to mediate the needs of two actors who would not have come together, or not in this form, without an independent broker. The brokerage platform operator takes on the requirements of a requester as well as the capabilities and offers of various providers. The platform operator provides the necessary coordination, mediation and selection process (request, offer and order) between the requester and a provider. Solutions, prices and available quantities are typically not fixed at the initial stage. Depending on the specific business model, the brokerage platform operator partly assumes a contractual commitment to the actors or just mediates the needs. Practical examples are V-Industry, Protiq and Spanflug.

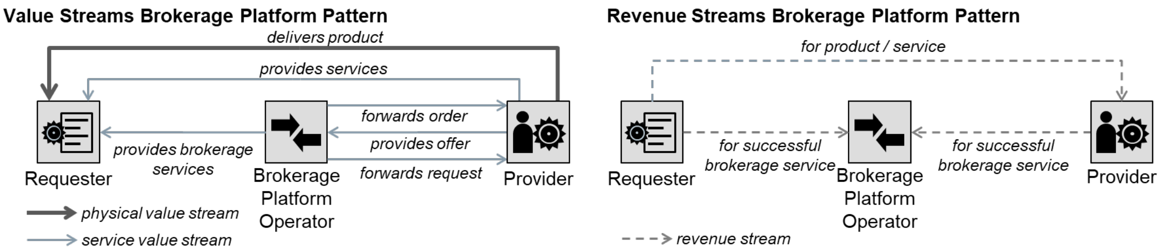


Figure 2: Value and revenue streams of the brokerage platform pattern

The brokerage platform provider offers the partners an independent mediation between two different actors, each with a different business interest, and simplifies the time-consuming coordination process from offer to order. Typically, the brokerage platform provider enables the easy establishment of contact and efficient transaction. The business partners benefit from the platform operator as follows:

- The requester gains the advantage of reducing the process of searching for suitable provider as well as the business coordination. The requester receives security for the transaction, as typically the platform operator provides a guarantee.
- The provider profits by an additional sales channel via the brokerage platform, where the customer acquisition is supported or partly already carried out by the brokerage of the platform operator.

The business roles involved in the value network receive revenue for the following value propositions:

- The brokerage platform operator receives a commission fee for each successfully brokered service or product, typically depending on the value of the transaction.
- The provider receives revenue from the sale of products and services.

4.3 E-Shop platform pattern

The purpose of the e-shop platform operator (shown in Figure 3) is to enable quick and easy buying and selling of products or services across different actors. The difference to the brokerage platform pattern is that an e-shop offers concrete services and products that are already announced with a fixed price and known availability. A direct purchase can be made via the e-shop platform without negotiations. The payment flows via the platform operator. Practical examples are Amazon for Business, Mercateo and Wucato.

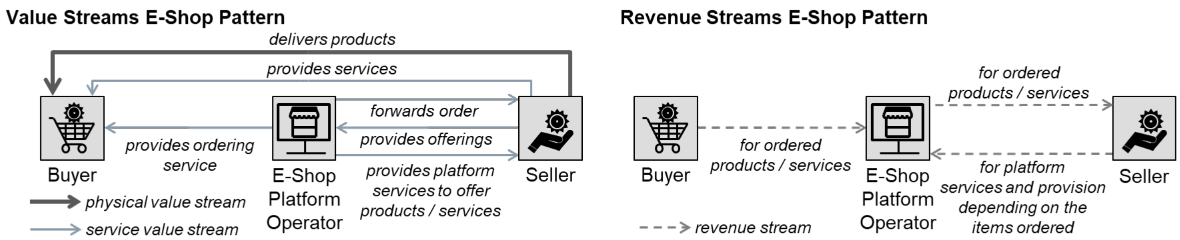


Figure 3: Value and revenue streams of the e-shop platform pattern

The e-shop platform operator offers a digital shop through which the purchase of products and services from many sellers and buyers can be handled centrally via the platform. The business partners benefit from the platform operator's offerings as follows:

- The seller can present and sell products and services to many customers with low effort via a new sales channel. In addition, processing tasks such as invoicing are taken over by the platform operator. However, the seller becomes transparent and comparable to competitors.
- The buyer benefits from the e-shop by being able to find various and comparable products centrally in one place and to order them uniformly. The buyer can purchase from different sellers in a short time and has only one central contact partner with the e-shop platform operator.

The business roles involved in the value network receive revenue for the following value propositions:

- The e-shop platform operator receives commission fees for the products and services sold via the e-shop. In some cases, the seller has to pay an additional one-time fee for the platform access.
- The seller receives revenue from the sale of products and services via the e-shop.

4.4 Supply Chain Management (SCM) platform pattern

The purpose of the SCM platform operator (shown in Figure 4) is to create a comprehensive exchange of information along the supply chain by bundling information and communication channels. For this purpose, an SCM platform offers companies the possibility to combine relevant information and process flows across companies and thus make information relevant for the delivery accessible in their supply chain (e.g. customers, suppliers). Practical examples are AirSupply, Siemens AX4 and Flexport.

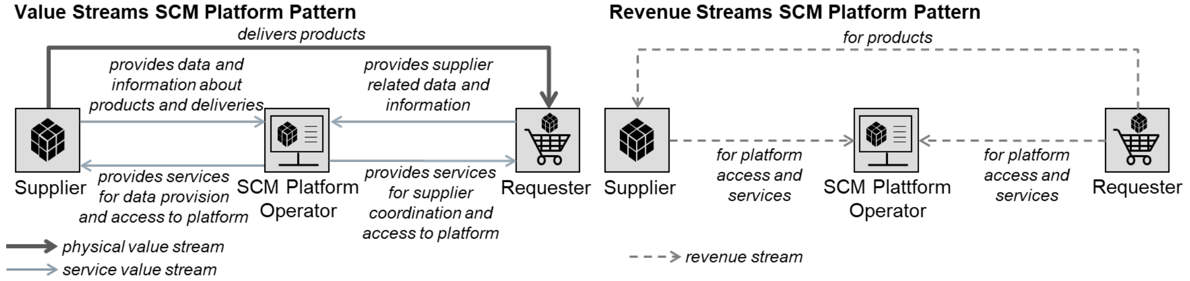


Figure 4: Value and revenue streams of the SCM platform pattern

An SCM platform provider offers an infrastructure so that information from different sources can be managed centrally and provided to partners. The SCM platform ensures access from different companies to uniform information (single point of truth). Additionally, the SCM platform provides software applications that optimise and automate process steps (e.g. sending invoices). The business partners benefit as follows:

- The requester and the supplier benefit from the SCM platform by receiving information on orders and deliveries uniformly via a single database. Furthermore, an optimisation of cross-company and internal processes as well as improved coordination with various supplier and requesters can be achieved, partly also through further services (e.g. delivery evaluations).

The business roles involved in the value network receive revenue for the following value propositions:

- The SCM platform operator receives revenue for the provision of the platform and additional services.
- The supplier receives revenues for the delivered products.

4.5 Content platform pattern

The purpose of the content platform operator (shown in Figure 5) is the uniform provision of information (e.g. image, text or price information) from different content providers via a single database. The content platform provides easy access to a uniform database with functions for analysis, searching or sorting. The information can be accessed simultaneously, but in contrast to the collaboration platform, they cannot be edited jointly by several users. In contrast to the SCM platform pattern, the content platform provides information not only for partners within a supply chain with the aim for optimizing supplier processes. Practical examples like Traceparts, Europages or E-Plan Data Portal provides data partly publicly accessible.

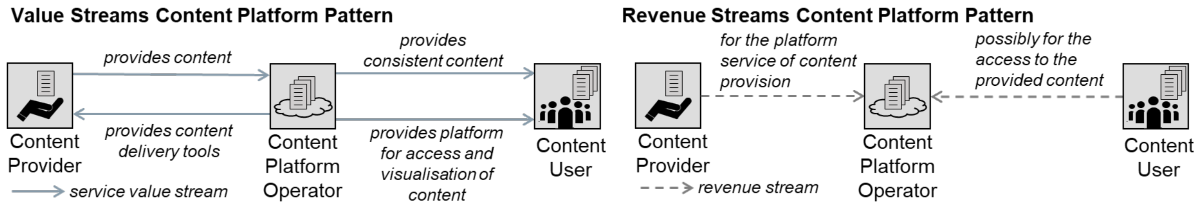


Figure 5: Value and revenue streams of the content platform pattern

The platform operator enables the partners with an easily accessible infrastructure for the uniform storage and provision of information. Thus, the operator provides the basis for a broad and comparable provision of information. The business partners benefit from the platform operator as follows:

- The content provider benefits by sharing information centrally to a broad user group in a uniform way (e.g. 3D-files). However, the standardization of the content by the content platform also makes the provider more comparable.

- The content user benefits by being able to access and use information from many providers on the content platform and reduces internal processes for searching and comparing information.

The business roles involved in the value network receive revenue for the following value propositions:

- The platform operator receives a fee from the content provider for access to the platform and partly for special placements of the content. In some cases, the platform operator also receives revenue from the content user for general access fees, access to premium content or further services.

4.6 Integrated platform pattern

The integrated platform pattern (see Figure 6) has the business purpose of providing a technical infrastructure as a basis for complementary solution offerings and mediating apps of 3rd parties via an app store. The complementary solutions are offered in relation or are executed on the integrated platform and the solution user can deploy further functions with low integration effort. Examples are Siemens Mendix or Amazon AWS.

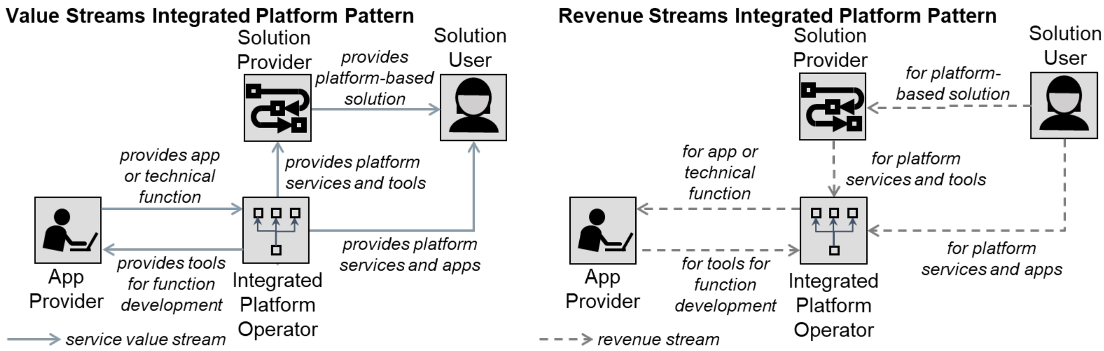


Figure 6: Value and revenue streams of the integrated platform pattern

The integrated platform operator offers an infrastructure on which other roles can offer further complementary solutions or expanding applications. Besides operation, the platform operator provides for example maintenance, security, and supporting services. The business partners benefit as follows:

- The solution provider is enabled to offer services or develop comprehensive products using features of the integrated platform or using provided apps by the app provider.
- The solution user benefits from the solution provided by the solution provider. The solution user can either use the solution without further adaptations, or can extend the solution via the integrated platform operator (e.g. via apps).
- The app developer benefits from the available development tools and can thus provide functions in addition to the integrated platform and sell apps via the integrated platform as a new sales channel.

The business roles involved in the value network receive revenue for the following value propositions:

- The integrated platform operator receives revenue from the solution provider respectively from the solution user for providing the platform (e.g. licence fee, monthly user fee). In addition, the platform operator receives income from the app provider for the provision of development tools, as well as commissions for applications successfully sold via the platform.
- The solution provider receives payment from the solution user for the platform-based solution.
- The app provider receives a usage-based fee for apps used by the solution provider or solution user.

4.7 Data exchange platform pattern

The data exchange platform owner (shown in figure 7) has the business purpose of providing a software solution for connectivity and data transmission between internal or external IT systems. The provided software solution is then integrated by a system integrator for specific uses. The data exchange platform operator typically uses the platform to exchange data internally, but maybe also with external partners. Examples for the data exchange platform pattern are KAMPF the@dvanced, Cybus and Ondeso.

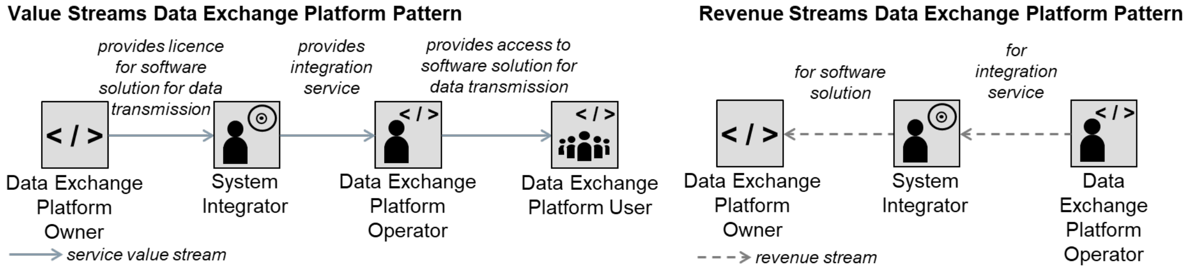


Figure 7: Value and revenue streams of the data exchange platform pattern

At this point, two roles need to be distinguished:

- The data exchange platform owner owns the intellectual property of the platform solution. However, the platform owner does not have any business relation with the end user or operate the solution, but provides a software solution that enables users to exchange data between different IT systems. Additionally, the platform owner provides updates and maintenance.
- The data exchange platform operator uses the software solution typically integrated and adapted for its specific IT environment by a system integrator to its users (typically within the own company).

The business partners benefit from the platform operator's offerings as follows:

- The data exchange platform user is technically enabled to exchange data between different IT systems.
- The system integrator takes on the role of an enabler by integrating and adapting the software to the data exchange platform operator.

The business roles involved in the value network receive revenue for the following value propositions:

- The data exchange platform owner receives revenue, typically in form of a licence model, for the provision and maybe for the usage of the software and, if applicable, for updates.
- The system integrator receives revenue for the integration service.

4.8 Collaboration platform pattern

The collaboration platform owner (shown in Figure 8) has the business purpose of providing a software solution for interdisciplinary and partly cross-company usage. For this purpose, the platform owner offers a software which can be integrated by the software integrator for specific usage, so that users can work simultaneously together (e.g. in a development project). The collaboration platform operator typically provides the integrated collaboration environment to internal and external company stakeholders. Practical examples are Siemens Teamcenter, COMOS and Contact Elements.

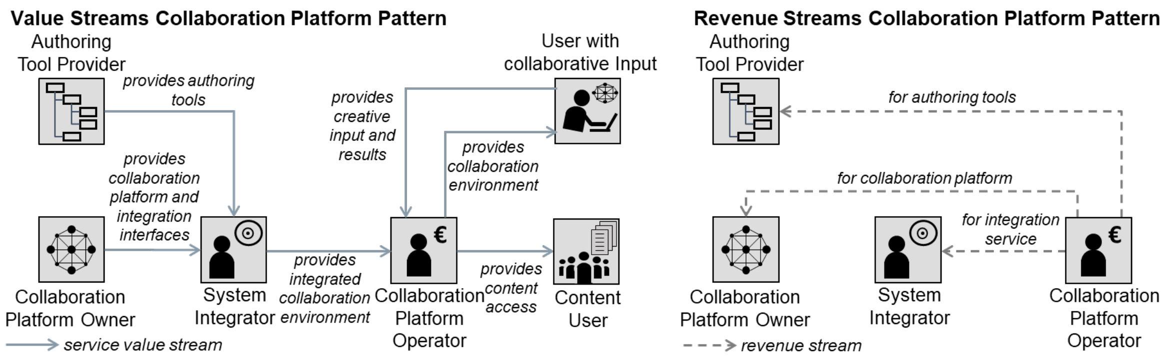


Figure 8: Value and revenue streams of the collaboration platform pattern

At this point, two roles need to be distinguished:

- The collaboration platform owner provides a software solution that enables users to work jointly on the same database. Typically, the collaboration platform owner provides updates and maintenance.
- The collaboration platform operator provides a specifically integrated collaboration environment, primarily to internal company stakeholders, but also to external company stakeholders, for collaborative working.

The business partners benefit from the platform operator's offerings as follows:

- The user with collaborative input benefits by being able to improve collaboration processes with internal and cross-company departments.
- The content user benefits by being able to easily access information from a unified database.
- The authoring tool provider acts as an enabler by providing additional software tools (e.g. CAD-Tools) necessary for collaboration.
- The system integrator enables the provision of an integrated collaboration environment of the collaboration platform operator.

The business roles involved in the value network receive revenue for the following value propositions:

- The collaboration platform owner receives revenue for the provision and maybe also for the usage of the collaboration software.
- The authoring tool provider receives revenue for providing the authoring tool.
- The system integrator receives revenues for the integration services.

5. Conclusion and Further Research

Based on a very heterogeneous range of different platform offerings in the manufacturing industry, it was possible to structure 160 identified platform offerings using a methodology concentrating on value and revenue streams from a business perspective. The research conducted shows that platform operators offer different value propositions to their partners, for example matching of actors, provision of new sales channel or enablement for further databased services. Thus, it can be concluded that platforms in the manufacturing industry also need to be differentiated from a business perspective. Despite the company-specific characteristics, eight patterns could be elaborated by abstracting the value creation relationships and revenue streams, which can be clearly distinguished from each other from a business perspective. Distinct means that a business decision is necessary to decide on a pattern. The focus on value propositions provides a decisive representation of what a platform operator offers partners in the value network. These business-relevant aspects are chosen as central criteria for the analysis and classification. Thus, a precise description and explanation of different platform patterns in the manufacturing industry could be achieved.

Further research is necessary to subdivide these general patterns into sub-patterns with more examples and to describe specific modes of transactions. These patterns form a well-grounded basis for future research in order to carry out differentiated analyses regarding dynamic effects, network effects and further success factors. Furthermore, further research can also examine what requirements companies should meet in order to implement the different platform patterns and how individual patterns can be combined. The aim is that in the future, industrial B2B platforms will become differentiated and transparent based on their concrete characteristics, and that specific success potentials can be identified in addition to B2C platforms.

References

[1] Kenney, M., Rouvinen, P., Seppälä, T., Zysman, J., 2019. Platforms and industrial change. Industry and Innovation 26 (8), 871–879.

[2] Pauli, T., Fielt, E., Matzner, M., 2021. Digital Industrial Platforms. Bus Inf Sys Eng 63 (2), 181–190.

[3] Evans, P.C., Gawer, A., 2016. The Rise of the Platform Enterprise. The Center of Global Enterprise. New York.

[4] Department of Economic and Social Affairs of the United Nations, 2008. International Standard industrial classification of all economic activities (ISIC), Rev. 4 ed. United Nations, New York.

[5] Kleinaltenkamp, M., Saab, S., 2009. Einführung in das Business-to-Business-Marketing, in: Kleinaltenkamp, M., Saab, S. (Eds.), Technischer Vertrieb. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–14.

[6] Falck, O., Koenen, J., 2020. Industrielle-Digitalwirtschaft-B2B-Plattformen. https://bdi.eu/publikation/news/Industrielle-Digitalwirtschaft-B2B-Plattformen. Accessed 28 March 2021.

[7] Gawer, A., 2014. Bridging differing perspectives on technological platforms. Research Policy 43 (7), 1239–1249.

[8] Parker, G., van Alstyne, M., Choudary, S.P., 2016. Platform Revolution. Norton & Company, New York.

[9] Engelhardt, S. von, Wangler, L., Wischmann, S., 2017. Eigenschaften und Erfolgsfaktoren digitaler Plattformen. https://www.digitale-technologien.de/DT/Redaktion/DE/Downloads/Publikation/autonomik-studie-digitale-plattformen.pdf?__blob=publicationFile&v=6. Accessed 28 March 2021.

[10] Täuscher, K., Laudien, S.M., 2018. Understanding platform business models: A mixed methods study of marketplaces. European Management Journal 36 (3), 319–329.

[11] Wortmann, F., Ellermann, K., Kühn, A., Dumitrescu, R., 2019. Typisierung und Strukturierung digitaler Plattformen im Kontext Business-to-Business, in: Vorausschau und Technologieplanung. 15. Symposium für Vorausschau und Technologieplanung, Berlin. 21. und 22. November 2019, 191–214.

[12] Plattform Industrie 4.0., 2019. Digitale Geschäftsmodelle für die Industrie 4.0. https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/digitale-geschaeftsmodelle-fuer-industrie-40.html. Accessed 28 March 2021.

[13] Chen, J., Cheng, Y., Chen, Z., Dorst, W., Gierging, L., Leonardy, H., Löwen, U., Zhao, H., 2020. Examples for Business Scenarios in Manufacturing Industry. https://www.gpqi.org/files/upload/china/studies/201015_GIZ_UseCases_BusinessScenarios.pdf. Accessed 28 March 2021.

[14] Plattform Industrie 4.0, Robot Revolution & Industrial IoT Initiative, 2021. Digital Platforms in Manufacturing Industries. https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/digital-platforms-in-manufacturing-2021.html. Accessed 28 Mai 2021.

[15] Bundesverband der Deutschen Industrie e.V. (BDI), 2020. Deutsche digitale B2B-Plattformen, https://bdi.eu/publikation/news/deutsche-digitale-b2b-plattformen/. Accessed 28 March 2021.

[16] Löwen, U., Hartner, F., 2020. Business model patterns for digital B2B platforms, in: 21. Leitkongress der Mess- und Automatisierungstechnik, Automation 2020, VDI Wissensforum, Baden-Baden, 741–754.

Biography

Fabian Hartner, M.Sc. (*1994) is a research associate and doctoral candidate at the Institute for Factory Automation and Production Systems (FAPS) at the Friedrich-Alexander University of Erlangen-Nürnberg (FAU) since 2019. He studied previously industrial engineering and management at FAU.

Prof. Dr. Ulrich Löwen (*1961) is working for Siemens as a subject matter expert on digitization in the manufacturing industry. He has many years of professional experience and is involved in various associations, consortia, and international collaborations. His work focuses on scenarios and use cases to illustrate the business value of new digital technologies.

Prof. Dr. Jörg Franke (*1964) is Head of Chair of the Institute for Factory Automation and Production Systems at the FAU. There he focuses research 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.

2nd Conference on Production Systems and Logistics

Sensitivity Analysis of a Simulation Model for the Determination of the Utilization of a Production Environment with the LPBF-Process

Aziz Ghannouchi[1], Tobias Stittgen[1], Johannes Henrich Schleifenbaum[1]
[1]RWTH Aachen University Digital Additive Production DAP, Germany

Abstract

Additive Manufacturing is on the brink of industrialization. In addition to the technical aspect, the aspect of production planning and optimization is continuously gaining importance with the technology being introduced into production environments. In order to allow a production logistics positioning for a production environment equipped with Laser Powder Bed Fusion machines, a simulation model has been developed at the RWTH Aachen University Chair for Digital Additive Production DAP. By means of this simulation model, production logistics key performance indicators (KPIs) such as the utilization are calculated as a function of nine parameters describing the production environment. To demonstrate the validity of the simulation model and allow a better understanding of its behaviour, the influence of each input on the utilization is investigated. In this study, a Global Sensitivity Analysis (GSA) based on the Monte Carlo method is performed to quantify the importance of the parameters. For this purpose, a Variance Based Sensitivity Analysis (VBSA) is conducted, followed by a Regional Sensitivity Analysis (RSA), and a GSA based on the PAWN method, allowing to perform all three settings of GSA, namely Factor Prioritization, Factor Fixing, and Factor Mapping. Finally, the knowledge brought by the sensitivity analysis allows for a user-friendly simplification of the production environment model, thus permitting a prioritized increase of the utilization.

Keywords

Monte Carlo Method; Latin Hypercube Sampling; Variance Based Sensitivity Analysis; Regional Sensitivity Analysis; PAWN; Additive Manufacturing; Laser Powder Bed Fusion; Production Logistics; LPBF

1. Introduction

Additive Manufacturing (AM) is on the brink of industrialization [1–3]. In particular, Laser Powder Bed Fusion (LPBF) is becoming more significant to manufacture individualized and complex parts [4]. As the demand in AM parts is growing, especially in the aviation, automotive and medical fields [5–8], AM service providing companies have to act efficiently from the operational point of view to be economically profitable. While from a technical aspect, LPBF is becoming mature and reliable, other aspects such as production planning are lagging behind. This is presented in the works of Stittgen, where the manufacturers using this technology are shown to only reach a limited utilization [1,9].

The overall objective of this study is therefore to generate insights regarding the production planning and data preparation for the LPBF process chain, data preparation being the process of assigning parts to so-called buildjobs. By using a Monte Carlo based Global Sensitivity Analysis (GSA), the importance of the factors describing the configuration of an LPBF production environment is examined, ultimately allowing

DOI: https://doi.org/10.15488/11243

to provide guidelines for industrial end users on how to increase the utilization of their capital intensive equipment.

2. Scope of the Study

First, the Dilemma of Operations Planning in the field of production logistics is described. Then, the simulation model of LPBF production environments developed at the RWTH Aachen Chair for Digital Additive Production (DAP) is presented. Finally, Sensitivity Analysis is introduced, along with the settings aimed for in this study and the methods used.

2.1 Dilemma of Operations Planning in Production Logistics

As almost every technical advantage in the field of manufacturing is exposed to continuous change, in order to be competitive on the market, manufacturing companies need to not only constantly improve their products and production processes, but also optimize aspects of production logistics. This means manufacturers must strive to pursue greater delivery capabilities and reliability with the lowest logistics and production costs possible [10]. Consequently, these companies must strive to minimize their throughput time, minimize their lateness, maximize the utilization of their machines and operators, and reduce their Work in Process (WIP). These four variables are defined by Nyhuis and Wiendahl in order to systematically describe and model the fundamental interrelations between production logistics Key Performance Indicators (KPIs) as shown in Table 1 [10].

Table 1: Definitions of the Production Logistics Key Performance Indicators

KPI	Definition
Lateness	The difference between the targeted processing end time and the actual processing end time
Throughput Time	The amount of time required for a production order from the completion of the previous order until the end of the considered process
Utilization	The ratio of the mean output to the possible maximum output
Work in Process	Cumulated work content (WC) of the orders queued for processing and those being processed

Nevertheless, when trying to maximize the utilization and minimize the three KPIs lateness, throughput time and WIP, the manufacturers encounter the so-called dilemma of operations planning [11]. This means that certain production logistics objectives, quantified by the KPIs, are in conflict [10]. Therefore, companies need to set goals on which production logistics KPIs need to be prioritized, and how to improve these prioritized indicators.

In this study, the focus is set on the utilization of the machines, as the equipment of an LPBF production environment is capital intensive and the KPI is therefore a priority from an internal point of view for the company.

2.2 Simulation Model of LPBF Production Environments

In addition to the technical challenges such as the development of stable process parameters, material considerations, or design guidelines for Additive Manufacturing, manufacturers using LPBF currently face operational challenges regarding the following aspects amongst others:

- Lot sizing: this aspect is fundamentally different from conventional processes, as LPBF allows to manufacture several different parts at the same time in one job.

- Machine and operator availability: as the operators are only required to mount and unmount the jobs. Their presence is not necessary while the jobs are running.
- The distribution of the work content as it does not follow a normal distribution and differs from conventional processes.

Compared to conventional manufacturing processes, LPBF production environments are characterized by a high level of automation, the non-necessity of tools, an unmanned production, and the possibility to manufacture parts of different geometries at the same time [12]. Furthermore, the work content distributions of LPBF jobs are longer than those of conventional processes such as CNC-milling [13]. Therefore, operations planning needs to be conducted specifically for this technology.

In this context, a simulation model following the described specifics has been developed at DAP. By means of this tool, an LPBF production environment is simulated, allowing to calculate the utilization of the machines while considering the inputs described below (a detailed description of the model can be found in [13]):

- N_M: the number of available machines in the production environment
- T_M: the necessary amount of time to mount the LPBF machines, measured in hours
- T_U: the necessary amount of time to unmount the LPBF machines, measured in hours
- O_1: the number of operators present in the first shift
- O_2: the number of operators present in the second shift
- O_3: the number of operators present in the third shift
- a_{WC}: Shape parameter of the distribution of the work content
- b_{WC}: Scale parameter of the distribution of the work content
- RR: Release rate, the rate of orders that are released for production, measured by production orders per day

Following the works of Stittgen, the work content of the production jobs is described following a Gamma-distribution, as opposed to the logistics operating curves models [9,13]. This distribution is chosen according to the field data gathered at industrial LPBF end users and is considered in the simulation model.

2.3 Sensitivity Analysis

As the production environment is defined in the DAP simulation model by using nine parameters, sensitivity analysis is a powerful tool to investigate the influence of the single factors on the KPI, allowing to identify the most influential ones, as well as the negligible ones. Sensitivity Analysis is defined by Saltelli as being the study of how the variation in the output of a model is attributed to variations of its inputs. The influence of the single input parameters on the outputs is therefore investigated [15,14]. In the context of this study, sensitivity analysis is used to investigate the influence of the parameters describing the configuration of a production environment on the utilization of the machines. In the following, the possible settings of sensitivity analysis are introduced, followed by the three methods used in this study.

2.3.1 Settings of Sensitivity Analysis

The goals pursued in the framework of a Sensitivity Analysis are called settings. The three following settings are distinguished:

Factor Prioritization allows to rank the parameters according to their influence on the variability of the output.

Factor Fixing allows to identify the parameters whose influence on the output is negligible.

Factor Mapping allows to determine the region of so-called behavioral inputs, which are inputs that lead to a chosen region of interest in the output. This region of interest is defined by a threshold.

2.3.2 Methods Used for this Study

In their work, Pianosi et al. list several GSA methods [16]. For this study, the following three methods are chosen, as they allow to to fulfill all three settings of Sensitivity Analysis. Furthermore, as they rely on three different mathematical principles, a cross-validation is performed to validate the results obtained.

Variance Based Sensitivity Analysis (VBSA) relies on the contribution of each parameter on the variance of the output. While the main effects measure the direct contribution of a parameter on the output, the total effects measure the impact of the studied parameter in interaction with all other factors. Therefore, main effects are suitable for the setting of Factor Prioritization, while total effects are used for the setting of Factor Fixing [15].

Regional Sensitivity Analysis (RSA) is aimed at identifying a region in the input space that leads to a chosen output region of interest. A threshold is chosen for the output to define the region of interest. The input sets are then split into behavioral and non-behavioral inputs according to the output they lead to, thus fulfilling the setting of Factor Mapping. Furthermore, the empirical cumulated density functions (CDFs) of the behavioral and non-behavioral inputs are defined and visualized. The maximum vertical difference between the two curves obtained for each parameter is calculated by using the Kolmogorov-Smirnov statistic, and represents a sensitivity index used for Factor Prioritization [17].

PAWN is a density-based method as it considers the cumulated density function of the output. The unconditional CDF and conditional CDFs of the output are defined, the conditional CDFs being the ones where individual factors are constrained within a chosen interval. The divergence between the CDFs is then calculated and represents the PAWN sensitivity indices, which fulfill the settings of Factor Prioritization and Factor Fixing [18,19].

2.3.3 Point Generation

The GSA methods described in the previous section rely on the repeated run of the simulation model on a large number of points. These points are generated with the aid of the Monte Carlo Method or a Quasi Monte Carlo Method [20,15]. In the works of Kucherenko et al., the Latin Hypercube Sampling (LHS) method is proven to be effective and leading to fast convergence [21].

3. Workflow of the Study

In this section, the steps of the conduction of the study are presented. First, the ranges and distributions of the input parameters are defined. Second, the input samples are generated. Then, the simulation model is run on the generated inputs. Afterwards, Global Sensitivity Analysis is performed by using the three methods introduced in the previous section. The results for all three settings are presented. In a further step, another GSA is performed on the behavioral inputs identified by means of Factor Mapping. Finally, relying on the generated data, further interpretation is presented regarding the planning of LPBF production environments.

3.1 Definition of Ranges and Distributions for the Input Parameters

The ranges and distributions of the parameters are defined as represented in Table 2. They are chosen such that feasible LPBF production scenarios are covered. The input parameters are assumed to be independent.

Due to the discrete nature of time steps used in the simulation model, the values of the mount and unmount time (T_M and T_U) are discrete.

Table 2: Ranges and Distributions of the Input Parameters of the GSA

Parameter	Distribution	Range	Unit
N_M	Uniform, discrete	[1 ... 10]	[-]
T_M	Uniform, discrete	[1 ... 4]	[h]
T_U	Uniform, discrete	[1 ... 4]	[h]
O_1	Uniform, discrete	[0 ... 5]	[-]
O_2	Uniform, discrete	[0 ... 5]	[-]
O_3	Uniform, discrete	[0 ... 5]	[-]
a_{WC}	See Equation 1	[1 ... 144]	[-]
b_{WC}	See Equation 2	[0.1 ... 25]	[-]
RR	Uniform	[0.1 ... 12]	Orders/Day

For the distribution parameters of the work content, the values follow customized distributions as shown in Equation 1 for a_{WC} (the shape parameter) and Equation 2 for b_{WC} (the scale parameter). The ranges are defined in the works of Stittgen [13]. The distributions represent realistic scenarios of LPBF buildjobs, as a uniform distribution leads to a large number of points with an unrealistically high mean work content.

$$x \in [1 ; 144] , PDF_{a_{WC}}(x) = \min\left(25, \frac{120h}{x}\right) \tag{1}$$

$$x \in [0.1 ; 25] , PDF_{b_{WC}}(x) = \min\left(144, \frac{120h}{x}\right) \tag{2}$$

3.2 Generation of the Inputs and Simulation Run

The input samples are generated using the Latin Hypercube Sampling method, each input representing a distinct production scenario [21]. N = 6,000 inputs are generated, covering the entirety of LPBF production scenarios. The investigated simulation model is then run on each of the generated inputs, allowing to calculate their respective utilizations. Because the model is complex and the conduction of the GSA requires a large number of model runs, the high performance cluster of RWTH Aachen University is used.

3.3 Global Sensitivity Analysis

In this section, the results of the Global Sensitivity Analysis are presented. In Figure 1, the sensitivity indices of each parameter obtained with all three methods are represented. All results are verified by means of a convergence analysis along with a robustness analysis [22].

3.3.1 Factor Prioritization

As described in Section 2.4, the setting of Factor Prioritization is fulfilled through all three methods: the parameters with the highest impact on the utilization of the machines are be deduced from the main effects of VBSA, as well as from the sensitivity indices delivered by RSA and PAWN. As shown in Figure 1 (a, c and d), the parameters with the highest sensitivity indices are the distribution parameters of the WC. As a result, they are the parameters with the highest impact on the utilization. The number of machines N_M as well as the production order release rate RR are the parameters with the subsequent sensitivity indices and are therefore likewise subsequent regarding the impact on the KPI.

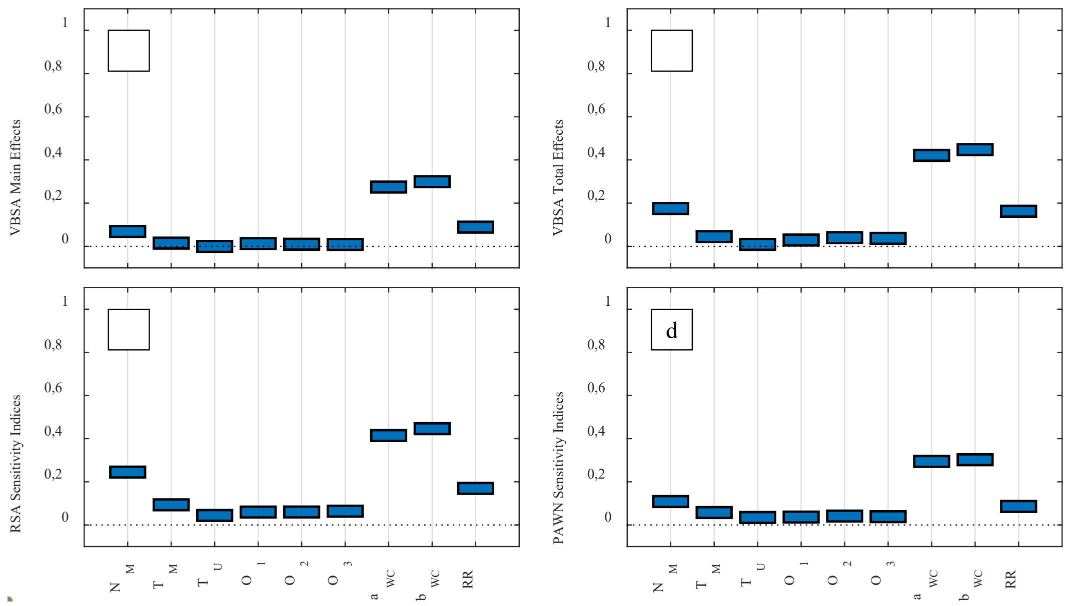


Figure 1: Sensitivity Indices of VBSA, RSA and PAWN for the Complete Design Space

3.3.2 Factor Fixing

As described in Section 2.4, the setting of Factor Fixing is fulfilled through the total effects of the VBSA as well as through the PAWN sensitivity indices. A threshold of 5% is chosen to identify a parameter as non-influential, meaning that a parameter with a sensitivity index below 5% is considered as non-influential on the utilization and can therefore be fixed to any value in order to simplify the model [22]. As shown in Figure 1 (b and d), the parameters T_M, T_U, O_1, O_2, and O_3 have a sensitivity index below the chosen threshold. Consequently, they have a negligible effect on the utilization of the machines of the production environment. The results of this setting are validated by using the method of Andres [23].

3.3.3 Factor Mapping

The setting of Factor Mapping is performed by means of the Regional Sensitivity Analysis (RSA). First, the output region of interest is defined by a threshold of 60%, this value being a benchmark observed with industrial end users and therefore represents a lower bound for the utilization that is to be aimed for by a company. This means inputs leading to a utilization greater than 60% are considered behavioral, while inputs leading to a utilization below this threshold are considered non-behavioral. The results of the setting of factor mapping are represented in Figure 2. The whole considered ranges of each parameter allow to reach the output region of interest. Therefore, for each value of each parameter, there are combinations of the other parameters that allow to reach a utilization of more than 60%. In addition, as the behavioral inputs are a set of points leading to a high utilization, a set of possible configurations for the LPBF production environment is obtained.

3.4 Sensitivity Analysis on the Behavioral Inputs

In the previous section, Global Sensitivity Analysis is performed on the whole design space of the inputs. In particular, by means of the Regional Sensitivity Analysis, a set of behavioral inputs leading to a utilization greater than 60% is identified. In this section, as suggested in the works of Noacco et al., in order to gain more insights on the behavior of the simulation model, a further GSA is performed after removing the poorly performing inputs [24]. Therefore a GSA is performed on those behavioral inputs. The aim is to investigate the importance of each of the parameters within the points that lead to a utilization greater than 60%.

As VBSA requires a tailored sampling, this method is not used for this investigation. Instead, the two methods used are Regional Sensitivity Analysis (RSA) and PAWN. As a further maximization of the utilization is aimed at, for the case of the RSA, a new threshold for the utilization is set to 90%. The results obtained by using these two methods are represented in Figure 3.

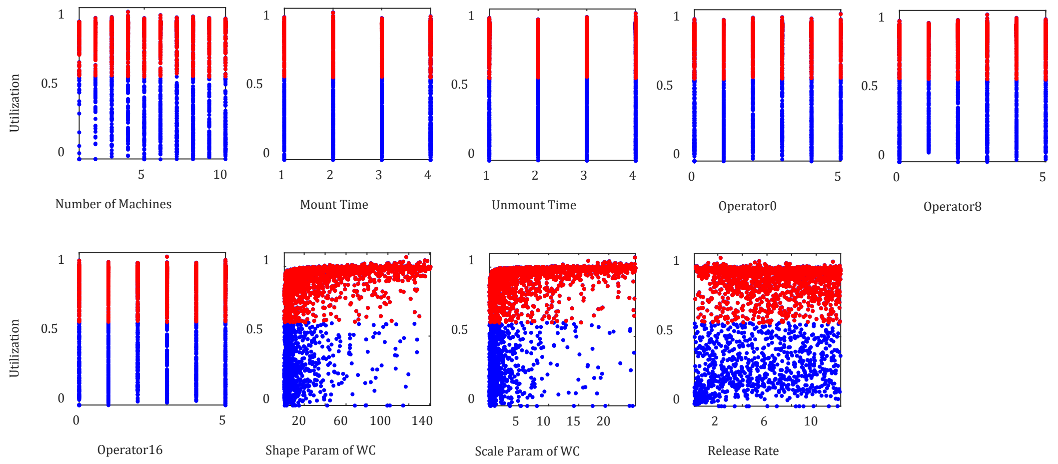


Figure 2: Scatterplots of the utilization as a function of all parameters

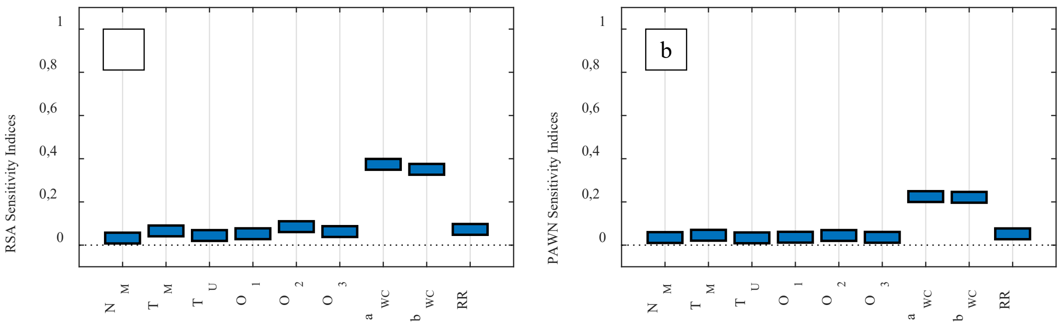


Figure 3: Sensitivity Indices of RSA and PAWN in the Set of Behavioral Inputs

As shown in Figure 3, and similarly to the analysis of the complete design space (Section 3.3), the parameters a_{WC} and b_{WC} have the highest sensitivity indices and are therefore the most influential parameters on the utilization of the production environment. As for the parameters T_M, T_U, O_1, O_2, and O_3, they have a low influence. Nevertheless, as opposed to the analysis in the complete design space, the parameters N_M as well as RR have a low impact within the set of behavioral inputs, as their sensitivity indices are below 5%. This means that once a utilization of more than 60% is reached, the number of machines in the production environment as well as the number of orders that are released per day have a lower impact on the utilization, meaning that a further maximization of the utilization is achieved only by considering the work content.

3.5 Further Interpretation of the Results

In this section, using the generated points and their corresponding outputs, further insights are presented regarding the planning of an LPBF production environment.

3.5.1 Impact of the Fluctuation of the Work Content on the Utilization

In this section, the impact of the fluctuation of the WC on the utilization is investigated. First, the variation coefficient of the WC is calculated for all inputs, as it is a standardized measure of dispersion. The points are then divided into equally spaced intervals according to the calculated value. Subsequently, for each interval, the mean utilization is calculated and represented as a function of the variation coefficient of the WC as shown in Figure 4. As presented in the figure, the mean utilization is beyond 90% for a variation coefficient lower than 0.1. The KPI decreases with an increasing variation coefficient until a convergence value close to 30% is reached. As a conclusion, for an LPBF production environment, in order to reach a high utilization, manufacturers need to aim for a low WC variation coefficient, meaning build jobs with a low fluctuation of the work content. The results are therefore compatible with the fourth law of production logistics introduced by Nyhuis and Wiendahl, stating that "the variance and the means of the work content determine the logistic potential of the shop" [10].

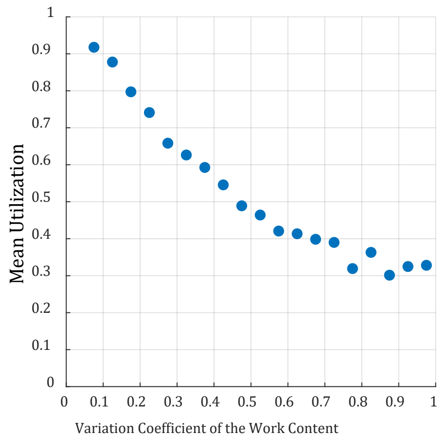


Figure 4: Mean Utilization as a Function of the Variation Coefficient of the WC

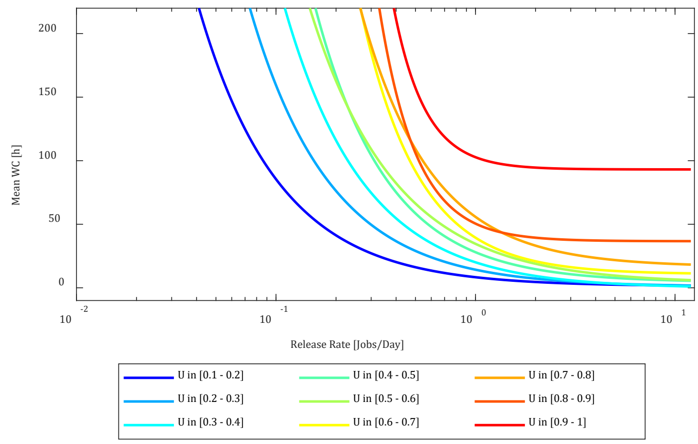


Figure 5: Mean Reachable Utilization as a Function of the Mean WC and the Production Order Release Rate

3.5.2 Guidelines for Data Preparation

A possible application of the results lies in the area of data preparation, a process step inherently connected to LPBF manufacturing and corresponding to the operations of lot sizing and nesting. As the number of operators as well as the mount and unmount time have a low impact on the utilization, a relationship is deduced between the mean work content, the production order release rate, the number of machines, and the utilization. First, the points are split according to the number of machines. Then, they are divided into ten equally spaced intervals according to their utilization (U), as shown in Figure 5 for the case of three machines. For each interval, the mean WC is plotted against the release rate. The proposed chart is therefore a tool used as a guideline for data preparation. For example, in the case of a targeted utilization of 80-90%, if a mean WC of 50 hours is sought after by the manufacturer, one production order has to be released per day. The knowledge of the target WC and the corresponding release rate allow a directed data preparation, so as to release jobs with the targeted length and frequency.

4. Conclusion and Outlook

In this study, a Global Sensitivity Analysis is conducted on a simulation model of an LPBF production environment. Performing the GSA leads to a better understanding of the behavior of the simulation model and of the impact of the single parameters of the utilization. The results obtained through the GSA allow to conclude the following regarding the utilization of an LPBF production environment:

- The most influential parameters are the distribution parameters of the work content, followed by the number of machines and the production order release rate.
- For the considered range of each input, there are combinations of the other parameters that allow to obtain a utilization of 60% or more.
- Beyond a utilization of 60%, while the distribution parameters of the work content are the most influential factors, the number of machines as well as the production order release rate have a low

impact on the utilization, meaning that they play a minor role to further maximize the production logistics KPI.

- A higher utilization of the machines can be reached with a low fluctuation of the work content.

In future works, further key performance indicators such as the throughput time need to be considered, since the utilization of the equipment is the only investigated KPI in this study. Furthermore, in the investigated model, the jobs are released according to the First In First Out (FIFO) principle. Therefore, in future works, an extension is to be developed in order to take other nesting and scheduling possibilities into consideration. Finally, further investigation is necessary to link the configuration of the production environments to factors such as lot sizing strategies, and the features of the parts to be manufactured within a given timeframe.

Acknowledgements

This work was carried out within the framework of the IDAM project (13N15084) funded by the German Federal Ministry for Education and Research (BMBF). Simulations were performed with computing resources granted by RWTH Aachen University under project rwth0638. The Global Sensitivity Analysis was performed by using the SAFE Toolbox, developed at the University of Bristol [24,25].

References

[1] Ciampi, E., Touloumian, A., 2018. Additive Manufacturing: Shaping up for Industrialization: Three catalysts are pushing 3D printing in a scaling-up phase. Oliver Wyman.

[2] Langefeld, B., Moehrle, M., Balzer, C., Schildbach, P., 2018. Advancements in metal 3D printing: Beyond powder bed - Additive Manufacturing on the brink of industrialization. Roland Berger.

[3] Proff, H., Staffen, A., 2019. Challenges of Additive Manufacturing: Why companies don't use Additive Manufacturing in serial production.

[4] Gebhardt, A., Hötter, J.-S., 2016. Additive manufacturing: 3D printing for prototyping and manufacturing. Hanser Publishers; Hanser Publications, Munich, Cincinnati.

[5] Javaid, M., Haleem, A., 2019. Current status and applications of additive manufacturing in dentistry: A literature-based review. Journal of oral biology and craniofacial research 9 (3), 179–185.

[6] Javaid, M., Haleem, A., 2018. Additive manufacturing applications in medical cases: A literature based review. Alexandria Journal of Medicine 54 (4), 411–422.

[7] Leal, R., Barreiros, F.M., Alves, L., Romeiro, F., Vasco, J.C., Santos, M., Marto, C., 2017. Additive manufacturing tooling for the automotive industry. Int J Adv Manuf Technol 92 (5-8), 1671–1676.

[8] Najmon, J.C., Raeisi, S., Tovar, A., 2019. Review of additive manufacturing technologies and applications in the aerospace industry. Additive Manufacturing for the Aerospace Industry.

[9] Stittgen, T., Schleifenbaum, J.H., 2019. Towards Operating Curves of Additive Manufacturing Equipment: Production Logistics and its Contribution to Increased Productivity and Reduced Throughput Time. 53rd CIRP Conference on Manufacturing Systems.

[10] Nyhuis, P., Wiendahl, H.-P., 2009. Fundamentals of Production Logistics. Springer Berlin Heidelberg, Berlin, Heidelberg.

[11] Gutenberg, E., 1951. Grundlagen der Betriebswirtschaftslehre: Erster Band: Die Produktion. Springer.

[12] Gibson, I., Rosen, D.W., Stucker, B., 2010. Additive Manufacturing Technologies. Springer US, Boston, MA.

[13] Stittgen, T., Schleifenbaum, J.H., 2021. Simulation of utilization for LPBF manufacturing systems. Prod. Eng. Res. Devel. 15 (1), 45–56.

[14] Saltelli, A., 2007. Sensitivity analysis in practice: A guide to assessing scientific models, Reprinted. ed. Wiley, Hoboken, NJ.

[15] Saltelli, A., 2008. Global sensitivity analysis. The Primer. John Wiley, Chichester England, Hoboken NJ.

[16] Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modelling & Software 79, 214–232.

[17] Spear, R.C. and Hornberger, G.M., 1980. Eutrophication in Peel Inlet - II Identification of Critical Uncertainties via Generalized Sensitivity Analysis.

[18] Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environmental Modelling & Software 67, 1–11.

[19] Pianosi, F., Wagener, T., 2018. Distribution-based sensitivity analysis from a generic input-output sample. Environmental Modelling & Software 108, 197–207.

[20] Lemieux, C., 2009. Monte Carlo and Quasi-Monte Carlo Sampling. Springer New York, New York, NY.

[21] Kucherenko, S., Albrecht, D., Saltelli, A., 2015. Exploring multi-dimensional spaces: A comparison of latin hypercube and quasi Monte Carlo Sampling Techniques.

[22] Sarrazin, F., Pianosi, F., Wagener, T., 2016. Global Sensitivity Analysis of environmental models: Convergence and validation. Environmental Modelling & Software 79, 135–152.

[23] Andres, T.H., 1997. Sampling methods and sensitivity analysis for large parameter sets. Journal of Statistical Computation and Simulation 57 (1-4), 77–110.

[24] Noacco, V., Sarrazin, F., Wagener, T., Pianosi, F., 2019. Matlab/R workflows to assess critical choices in Global Sensitivity Analysis using the SAFE toolbox. Earth-Science Reviews 194, 1–18.

[25] Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. Environmental Modelling & Software 70, 80–85.

Biography

Aziz Ghannouchi is a mechanical engineering post-graduate student at RWTH Aachen University. Since 2019, he has been a student assistant at the Digital Production group of the RWTH Chair of Digital Additive Production DAP.

Tobias Stittgen studied Industrial Engineering with a focus on production technology at RWTH Aachen University. After graduating, he was the responsible sales engineer for European automotive OEMs and suppliers at Laserline GmbH. Having returned to Aachen in 2017, he is now leading the Digital Production group at the RWTH Aachen University Chair Digital Additive Production DAP and at the same time responsible for sales and technology at the Aachen Center for Additive Manufacturing ACAM.

2nd Conference on Production Systems and Logistics

Identification of Cyber Security Risks in Subscription-based Business Models for Manufacturing Companies and Derivation of Suitable Measures

Günther Schuh[1], Volker Stich[1], Jan Hicking[1], Lars Kaminski[1], Jacques Engländer[1], Anna Majchrzak[1]

[1]*Institute for Industrial Management, FIR at RWTH Aachen University, Campus-Boulevard 55, Aachen 52074, Germany*

Abstract

In the age of digitalization, manufacturing companies are under increased pressure to change due to product complexity, growing customer requirements and digital business models. The increasing digitization of processes and products is opening up numerous opportunities for mechanical engineering companies to exploit the resulting potential for value creation. Subscription business is a new form of business model in the mechanical engineering industry, which aims to continuously increase customer benefit to align the interests of both companies and customers. Characterized by a permanent data exchange, databased learning about customer behavior, and the transfer into continuous innovations to increase customer value, subscription business helps to make Industry 4.0 profitable. The fact that machines and plants are connected to the internet and exchange large amounts of data results in critical information security risks. In addition, the loss of knowledge and control, data misuse and espionage, as well as the manipulation of transaction or production data in the context of subscription transactions are particularly high risks. Complementary to direct and obvious consequences such as loss of production, the attacks are increasingly shifting to non-transparent and creeping impairments of production or product quality, which are only apparent at a late stage, or the influencing of payment flows. A transparent presentation of possible risks and their scope, as well as their interrelationships, does not exist. This paper shows a research approach in which the structure of subscription models and their different manifestations based on their risks and vulnerabilities are characterized. This allows suitable cyber security measures to be taken at an early stage. From this basis, companies can secure existing or planned subscription business models and thus strengthen the trust of business partners and customers.

Keywords

Cyber Security, Subscription Business, Manufacturing Companies, Digital Transformation

1. Introduction

1.1 Initial situation

National initiatives such as the National Climate Protection Plan 2050 are driving the goal to achieve high standards of economic, social, or environmental sustainability. Individual components that contribute to increasing sustainability, such as the circular economy, are therefore highly relevant to all social sectors in Germany [1]. Enterprises in the field of mechanical and plant engineering have recognized this fact and are setting goals for the sustainable orientation of their corporate activity [2], since a long time of use has a major impact on CO2-emissions generated during the product life cycle of machinery and equipment [3]. This, however, contradicts the prevailing business model in the manufacturing industry - the transactional business model, which consists of the one-off sale of products. Therefore, the manufacturer faces the challenge of maximizing two conflicting goals, i.e., increasing the useful life of its machines and, at the same time, maintaining profit through one-time sales.

As trends, both a circular economy and digitalization have a significant impact on markets and core elements of enterprises' business models. It is apparent that the mere sale of goods, i.e. the core business, delivers a significantly lower contribution margin than services and after-sales services in the enterprises of the machinery and plant-engineering sector [4]. The shift from manufacturing to software- and data-driven services shape the view of how and for what customers pay. This leads to a turn from revenue to user fees. The convergence of software and hardware components acts as a catalyst and foundation for these developments. The associated networking and resulting databased services open up considerable potential for digitization. Digital business models such as subscriptions are ultimately the logical consequence of this development. Numerous machine and plant manufacturers are already developing corresponding solutions today. Moreover, in current times of crisis, it is evident that those enterprises that pursue subscription business models are performing better than other enterprises in navigating the COVID 19 crisis. Their revenue grew by 9.5% in the first quarter of 2020 whereas the S&P 500 only posted 1.9% growth [3].

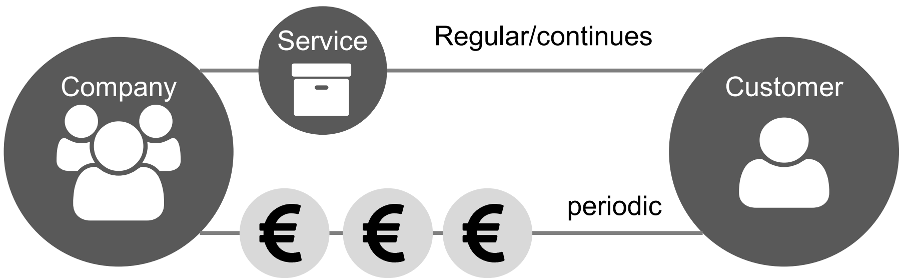


Figure 1: Subscription business model scheme

An enterprise enters a long-term relationship with its customers by providing services. Since this relationship has a great impact on the output and performance of manufacturing companies, it thrives on the mutual trust relationship between the partners involved [1]. Currently, the lack of trust in the context of cross-company data flows is a major obstacle to the implementation of new business models [2]. The increasing number of partners in a value network requires joint responsibility.

Besides, databased services and digital business models, such as subscription, require cross-linking of machinery. Existing machines (legacy) and modern machines are considered separately for this purpose. First, legacy machines are retrofitted, whereby the retrofitted solution must be guaranteed concerning cyber security. Even modern machines are not safe per se, even if they are based on existing standards (e.g., OPC-UA) and norms (e.g., IEC 62443). After all, only properly configured solutions offer the best possible protection. The integration of machines into the customer's network and transmission of the data to the machine and plant manufacturer are complex questions that have yet to be answered.

Digital transformation means that enterprises need to increasingly provide proof of trust to participate in value creation networks and to remain competitive [4]. Thus, appropriate cyber security will become a success factor for machine and plant manufacturers and manufacturing companies with digital business models in the future. A lack of expertise in the IT area and a lack of assistance, especially for small and medium-sized enterprises (SMEs) on which this paper will focus, prevent an efficient approach. Solution approaches, especially for SMEs, need an appropriate level of complexity, necessary investment, and applicability to enable efficient protection. To develop their business models and help shape their markets, the need for tried-and-tested solution patterns for their technical design is a necessity.

Software companies such as Microsoft with its Office 365 product and production enterprises such as the car manufacturer Volvo with its "Auto-Abo" demonstrate that subscription business models will increasingly prevail despite the aforementioned challenges. It is no longer sufficient to simply ensure IT security of individual components; instead, the need for action at the level of the overall system is necessary. This ensures that communication paths and corresponding technologies are designed to be as secure as possible across the entire value chain from the provider to the customer.

The research presented in this paper addresses the above issues by proposing the examination of subscription business models based on constituent characteristics from a cyber security perspective. It describes an approach to investigate the resulting attack vectors and their interaction with resulting measures. This perspective intends to sensitize subscription service providers as well as their customers to the relevance of cyber security and to contribute to accelerating acceptance and implementation in SMEs.

1.2 Structure of the paper

Firstly, basic terminology is explained to provide a general basis of understanding regarding the introduced issues. This is followed by an explanation of the potential and relevance of subscription business models and a more detailed description of the specific challenges involved. Subsequently, we describe existing management and security frameworks, as well as their focus on the described topic. On this basis, we present a systematic approach to solving the challenges described at the beginning. This approach is divided into four parts. First, we present different specifications of subscription business models, whereupon possible attack vectors are derived. Based on these attack vectors, an approach for determining suitable measures for these attack vectors is presented to subsequently describe them concerning their interactions. Finally, we highlight how to classify the findings from the previous steps using thread modeling.

2. State of the art

2.1 Terminology

Cyber security

Cyber security deals with protecting computers, servers, mobile devices, electronic systems, networks, and data from malicious attacks. It is also referred to as information technology security or electronic information security [5]. Additionally, it deals with various fields of application such as network security, application security, information security, disaster recovery, and business continuity, and end-user education.

Digital business ecosystems

From a business point of view, “[...]ecosystems encompass a set of actors that contribute to the focal offer’s user value proposition” [6]. A digital ecosystem is an assemblage of interconnected information technology resources and diverse stakeholders that can function as a cohesive unit. These include customers, research and collaboration partners, suppliers, external service providers, software and hardware, as well as third-party data service providers [7]. As such, it is understood as an environment of digital platforms that enable enterprises to position their products and services in the digital world [8].

Digital business models

A digital business model is a form of value creation that is based on developing added value and benefits for customers utilizing innovative digital technologies. Such a business model aims to generate a significant advantage for which customers are willing to pay [9]. As such, it is understood as a set of rules created by a company to organize its business processes in the most profitable way possible. It includes all aspects relevant to value creation, such as the respective range of products and services, addressed customers and target groups, how customer communication is performed, how services are rendered, revenue is generated, and transactions are carried out. All these aspects generate information that is collected, processed, analyzed, or communicated further using digital technologies. As a result, all processes are automated, and process chains are better coordinated. In this way, digital technologies streamline process chains, increase efficiency and maintain a company's competitiveness [10]. The transformation to digital business models is shifting the focus of attention towards customers, making it increasingly important to collect their data in the form of user profiles to adapt products and services more closely to their wishes and needs and even to anticipate them.

Threat Modeling

A threat model is a structured representation of all the information affecting the security of an application. Threat modeling is a process for capturing, organizing, and analyzing all information relevant to cyber security. As a result, it enables an informed decision-making process regarding the security risk of an application. Also, a priority list of cyber security enhancements is created for the concept, requirements, design, or implementation of each application. By identifying targets and vulnerabilities and then defining countermeasures to prevent threats or reduce their impact, the security of the system is holistically improved. [11]

2.2 Norms, standards and methods

Cyber security is referred to as a moving target in constant motion, 100% protection cannot be achieved. New measures are implemented with a certain delay and in an insufficient form. Moreover, the relevance to specific applications is often not sufficiently considered. As a result, there exist numerous frameworks with a different focus for securing cyber security like ISO IEC 27001:2015, NIST, or the IEC 62443, which are either too extensive, too complex, or not designed for the needs of companies.

ISO IEC 27001

The ISO IEC 27001 as an international and cross-sector standard supports companies to manage their information security. This standard contains requirements and specifications for the implementation, maintenance, and continuous improvement of information management systems (ISMS). Besides, methods are described to make the handling of risks related to information security controllable. The presentation of this information is at a generic level to ensure the broadest possible applicability to all stakeholders. [12]

IEC62443

The international series of standards IEC 62443 defines standards and guidelines for the cyber security of "Industrial Automation and Control Systems" (IACS) and provides basic guidelines for operators, integrators, and manufacturers concerning the design, implementation, management, manufacture, and operation of IACS [13]. The core areas are general principles, safety requirements for operators and service providers, safety requirements for automation systems, and safety requirements for automation components. The general principles describe, for example, basic concepts such as defense-in-depth or even fundamental requirements and refer to other parts of the standard for concrete implementation. Guidelines for the implementation of organizational measures are contained in security requirements for operators and service providers. It provides technical aspects such as security level and security requirements for automation systems. The fourth area aims specifically at the product and component view (sensors, interfaces, chips, etc.) and focuses more on manufacturers. [14]

IDS - An approach to securing digital ecosystems

The International Data Spaces (IDS) initiative aims to create a secure data space that enables enterprises of all sizes and from different industries to manage their data assets in a self-sufficient way. The IDS reference architecture model includes all components required for secure exchange and combination of data in ecosystems. The overall architecture consists of four sub-architectures: business architecture, software architecture, security architecture, and data and service architecture. [15] The IDS are distributed networks of data endpoints, provide standardized interfaces and define technical terms of use for data. In addition to legal and organizational contractual rules, it is also possible to formulate and implement technical terms of use. This enables data providers to retain sovereignty over their data when exchanging it. [16]

2.3 Potentials of subscription business models

Subscription business models are characterized by a consistent focus on customer benefits. The aim is no longer to sell the customer individual products or services, but rather to offer access to a constantly improving service. Thus, the subscription business changes the supplier-customer relationship in many respects. In the traditional transaction business, everything was geared to the sale and conclusion of high volumes. In the subscription business, the nature of the business relationship changes fundamentally. Due to the continuous data-technical connection to the customer and thus the insight into the use of the service offers in the field, the provider now has the opportunity for continuous, customer-specific service improvement. This enables customer and provider to enter into a participatory business relationship. We assume that both benefit directly in monetary terms from successful use or an increase in the customer's productivity. HARLAND AND WENGER characterize subscription businesses based on the following four constituent features along the business model dimensions described by GASSMANN ET AL. [17–19]:

1. Revenue mechanics: A subscription is characterized by periodic, unit-of-performance payment streams.
2. Value proposition: A subscription strives for the continuous performance increase of the customer benefit. To this end, the customer benefit is quantified.
3. Value architecture: Subscription is based on knowledge of changes in individual customer benefits. This requires integrating service bundles with customers, improving service delivery through continuous releases, and analyzing individual customer user data based on a customer ID.
4. Customer: A subscription describes a long-term, collaborative partnership. For this purpose, the target systems of provider and customer must be harmonized.

2.4 Research gap

Although many enterprise security frameworks exist, as can be seen in the state of the art section, they are either designed for universal application and thus often too broad for SMEs, or they take only a general perspective without directly addressing the security needs of digital business models. The paradigm shift from business models to collaborative value networks with active data exchange requires a new perspective from an information security point of view. Subscription business models, in particular, are characteristically predestined to represent this change since they are based on highly frequent, condensed data exchange, a high degree of dependence of the partners on each other, and trust. Therefore, an investigation of information security in subscription models needs a target-oriented approach to conclude the security requirements of new digital business models.

3. Research Results

3.1 Identification of cyber security challenges in subscription business models

The motivation and structure of subscription business models described above suggest the need for intensive cyber security consideration. Figure 2 illustrates the various attack vectors that result from the way subscription business models work.

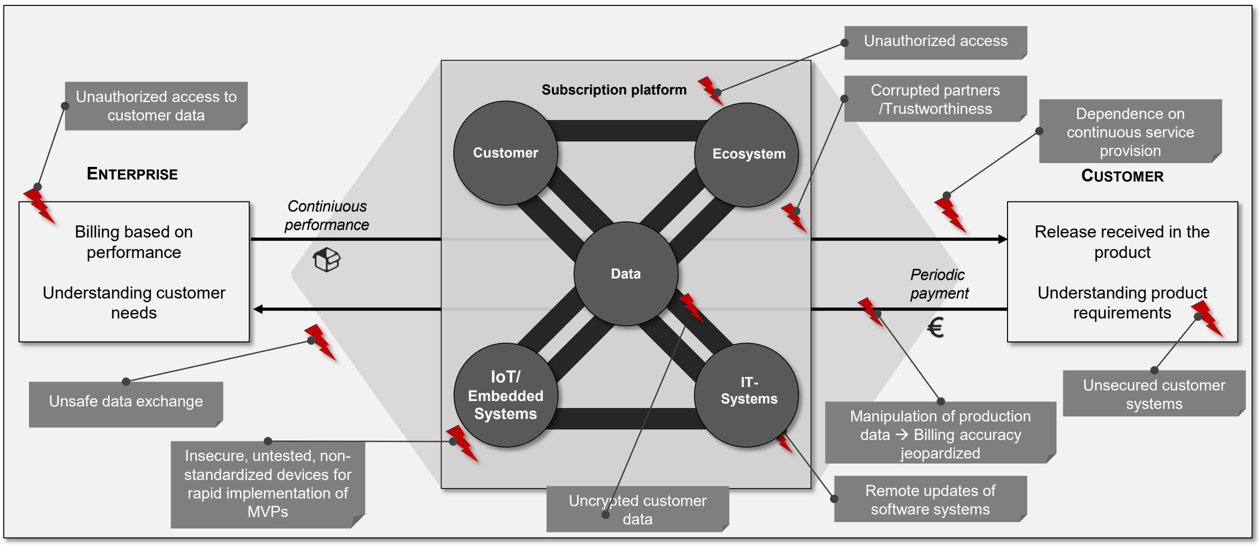


Figure 2: Attack vectors on subscription business models

The features presented in section 2.3 result in properties that require end-to-end protection. To ensure continuous service provision, we examine the dependencies on other systems and the interdependencies between the Stakeholders. In complex structures, continuous service provision is a critical factor for success on the customer side. As a result, this service provision is inevitably guaranteed constantly. It is also important to ensure correct billing depending on the type of subscription. This is done based on the time of usage, the number of units produced, or purely on a performance basis. The integrity and correctness of the data play a critical role. To bring a continuous increase in customer use in the sense of the value proposition, frequent exchange of, for example, production data, and performance data are necessary. This data is primarily used to optimize the product or the processes. On the buyer's end, there is a fear of intrusion into the protected corporate network and thereby endangering or losing sensitive product and production data. For this reason, enterprises prevent communication in and out of their network. In cases where this is inevitable (e.g., condition monitoring or remote maintenance), extensive legal safeguards are implemented. On the vendor or supplier end, the consumption volumes and correctness of usage information are indispensable to falsify payment information and save costs. Although buyers are merely users of the product, they have far-reaching access and thus also manipulation options. These options include data manipulation through unauthorized access to IT systems. This compromises the accuracy of production data and the associated invoice data, resulting in significant monetary damage. Since the financial damage caused by incorrect usage information or faulty configuration is substantial, it is in the vendor's best interest to secure the product against unauthorized manipulation by the customer or by third parties. This results in a paradigm shift in the area of facility operation: neither the physical access to the facility nor the IT network in which the facility is located is considered trustworthy anymore. The close relationship and connection between provider and customer make both sides potential gateways for hackers who can significantly influence the other party. It is becoming apparent that the interests of customers and providers are moving closer together. Information security must therefore shift from a purely singular view of one's own company to a holistic view as an ecosystem.

3.2 Top-down approach

In the following, we explain the schematic approach of the method developed in the future and the individual steps briefly. We divided the procedure for achieving a suitable method into four phases (Figure 3). First, we describe subscription business models based on their constituent characteristics from a cyber security perspective to derive use cases and resulting attack vectors based on these requirements. For this purpose, we assign subscription models to different types, based on STOPPEL supplemented by cyber security relevant aspects [20].

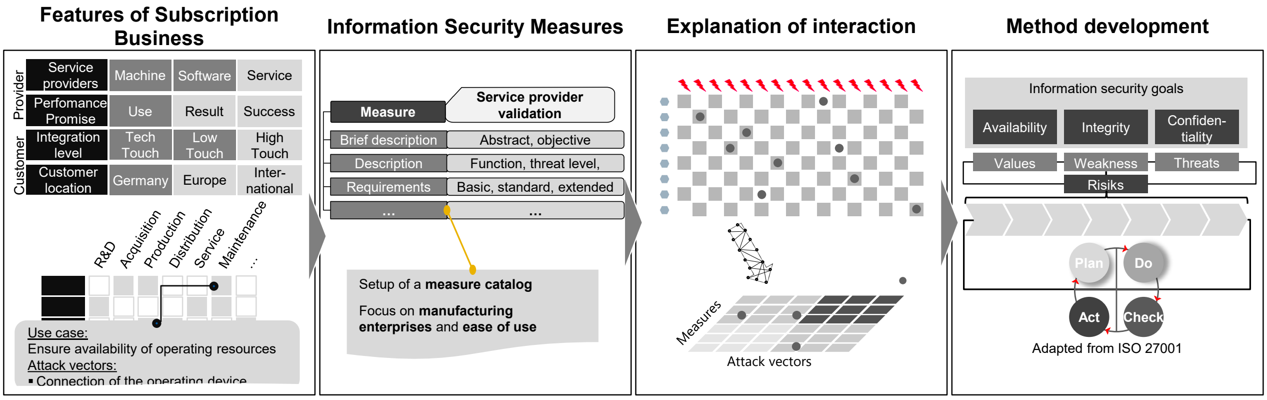


Figure 3: Components of the method development

The first type describes the pure availability of a product and represents the classic subscription model. At this stage, the supplier provides the customer with the machine or its service provision and maintenance in exchange for a periodic continuous payment. In addition to the transmission of payment data, there is no extensive data storage or transmission. Possible attack vectors on this type are, for example, manipulation of the payment information or influencing the machine at the customer's site, physically or via digital interfaces, to interrupt service provision.

The second type, usage-based, refers to billing based on effective production time or usage tracked in the payment interval. Terms such as pay-per-use or pay-per-hour are common in this context [21]. To enable this type of subscription, highly accurate tracking of machine usage is required. This is often measured in terms of uptime, to enable error-free and unambiguous billing. There is a risk that the transmitted usage time is manipulated and thus the billing does not correspond to the performed usage time. This can happen both by manipulating the machine itself, or the responsible functions and sensors and by manipulating the data itself.

This is followed by the third, result-dependent type. The reference figure used is the result of the use of infrastructure. Such systems are known as pay-per-unit or pay-per-part solutions, especially in operator models [22]. Whereas in usage-based models, the primary interest of the provider lies in the most intensive use of the infrastructure by the customer. In contrast, in outcome-based pricing systems, the provider focuses on achieving the outcome as efficiently as possible, according to which the provider is paid and for which he must pay. A possible attack vector is represented by the manipulation of the provided environment. There is the risk that the efficiency of the environment, as well as the quality and quantity of production, is negatively influenced, resulting in the need to prevent manipulation.

We describe the last type as performance-based. The reference of the value of the systems is the economic success resulting from interactions between supplier and customer. Customers pay a price for economic success, which is measured, for example, in terms of cost reductions, higher contribution margins, or profits. With a performance-based system, the provider also partially assumes the value risk in addition to the risks already discussed. The customer only has to pay if economic success is generated at all from the interaction with the provider. If this success is lower, the customer also pays a lower price. Preventing value creation or optimization through hacker attacks, denial of service or other attacks are critical variables to consider here, which greatly increase the risk on the provider's side.

In the second phase, after the features of the respective subscription model characteristics and their possible attack vectors have been discussed, we identify measures that minimize the described attack surface. To this end, we describe measures and their precise characteristics and scope concerning one or more attack vectors and classified them based on an evaluation catalog yet to be developed. This evaluation and description, which is still being prepared, is based on KÖNIGS [23]. Subsequently, in phase three, we explain cause-effect relationships between the identified attack vectors and the cataloged measures. We develop influence diagrams and influence matrices, based on the approach by PROBST AND GOMEZ [24], that make the relevance of the relationships for users transparent and understandable.

Based on the fundamentals now obtained, a method (phase 4) will be designed that secures the individual characteristics of subscription business models. To achieve this goal, we combine existing management approaches such as the ISO 27001 and the ISIS12 [25] procedure with a procedure based on the threat modeling described above. This method is composed of four core components:

1. Modeling of the targeted system to be built or modified based on characteristic properties
2. Use existing models such as STRIDE [26] to identify threats to the system described in step one
3. Addressing the identified threats based on the model applied in step two
4. Validation of completeness and efficiency

This approach of threat modeling serves to design methods and frameworks that are as close to the application as possible and that can efficiently deal with specific cyber security challenges. The intended result is a method that allows determining the own expression of the subscription model including the resulting attack vectors, corresponding measures, and their interaction to ensure a deployment roadmap for the implementation of the measures and transparency.

4. Contribution and Discussion

The contribution of this paper is a new way of looking at digital business models, especially subscription business models, from a cyber security perspective. We highlighted the motivation and necessity of these business models and described the resulting challenges. Furthermore, we explained the constitutive features of subscription transactions in more detail and, based on this, described different types and their characteristics. On this basis, specific cyber security challenges were described and their potential impact highlighted. In doing so, we have emphasized the criticality of the need for more intensive consideration. Finally, we presented an approach on how we want to solve this issue methodically. To this end, we outlined a four-step approach. The components will be detailed in the future and molded into a methodological procedure. The overall goal is to secure provider competitiveness on the one hand, and on the other hand, how customers are facilitated in their decision to use and accept subscription business models.

References

[1] Bundesministerium für Wirtschaft und Energie (BMWi), 2020. IIoT Value Chain Security – The Role of Trustworthiness, Berlin.

[2] Lassnig, M., Stabauer, P., Breitfuß, G., Müller, J.M., 2019. Erfolgreiche Konzepte und Handlungsempfehlungen für digitale Geschäftsmodellinnovationen, in: Meinhardt, S., Pflaum, A. (Eds.), Digitale Geschäftsmodelle – Band 1. Geschäftsmodell-Innovationen, digitale Transformation, digitale Plattformen, Internet der Dinge und Industrie 4.0. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 201–219.

[3] Widjaja, R., 2020. Focus: Subscription Economy: The Future of Business Models. https://cigp.com/insights/focus-subscription-economy-the-future-of-business-models. Accessed 12 February 2021.

[4] Henke, M., Schulte, A.T., Jakob, S., 2020. Blockchain-basiertes Supply Chain Management, in: Hompel, M. ten, Bauernhansl, T., Vogel-Heuser, B. (Eds.), Handbuch Industrie 4.0. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 599–615.

[5] Kaspersky, 2021. What is Cyber Security? https://www.kaspersky.com/resource-center/definitions/what-is-cyber-security. Accessed 9 March 2021.

[6] Kapoor, R., 2018. Ecosystems: broadening the locus of value creation. J Org Design 7 (1).

[7] Brush, K., 2019. Definition Digital Ecosystem. https://searchcio.techtarget.com/definition/digital-ecosystem. Accessed 9 March 2021.

[8] Canteli, A., 2018. Digital business ecosystem. OpenKM. https://www.openkm.com/blog/digital-business-ecosystem.html#:~:text=Written%20by%20Ana%20Canteli%20on,into%20clients%20and%2For%20subscribers. Accessed 18 March 2021.

[9] innolytics.ag, 2020. What is a digital business model? https://innolytics-innovation.com/digital-business-model/#:~:text=A%20digital%20business%20model%20is,customers%20are%20willing%20to%20pay. Accessed 18 March 2021.

[10] Bundesministerium für Wirtschaft und Energie (BMWi), 2017. Digitale Geschäftsmodelle: Themenheft Mittelstand-Digital, Berlin. https://www.bmwi.de/Redaktion/DE/Publikationen/Mittelstand/mittelstand-digital-digitale-geschaeftsmodelle.pdf?__blob=publicationFile&v=16.

[11] OWASP Foundation, 2021. Threat Modeling. https://owasp.org/www-community/Threat_Modeling. Accessed 23 March 2021.

[12] TÜV Süd, 2021. ISO/IEC 27001 - ISMS-Zertifizierug. https://www.tuvsud.com/de-de/dienstleistungen/auditierung-und-zertifizierung/cyber-security-zertifizierung/iso-27001. Accessed 27 March 2021.

[13] Deutsche Kommission Elektrotechnik Elektronik Informationstechnik in DIN und VDE (DKE), 2020. IEC 62443: Die internationale Normenreihe für Cybersecurity in der Industrieautomatisierung. https://www.dke.de/de/arbeitsfelder/industry/iec-62443-cybersecurity-industrieautomatisierung. Accessed 22.02.21.

[14] Seipel, C., 2020. EC 62443: Die internationale Normenreihe für Cybersecurity in der Industrieautomatisierung. DKE. https://www.dke.de/de/arbeitsfelder/industry/iec-62443-cybersecurity-industrieautomatisierung. Accessed 19 March 2021.

[15] Otto, B., Jürjens, J., Schon, J., Auer, S., Menz, N., Wenzel, S., Cirullies, J., 2016. Industrial Data Space - Digitale Souveränität über Daten (White Paper). White Paper, München, 40 pp.

[16] Forschungszentrum Data Spaces, 2020. Allgemeine Fragen zu den International Data Spaces. Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. https://www.dataspaces.fraunhofer.de/de/faq.html. Accessed 23 March 2021.

[17] Ebi, M., Hille, M., Doelle, C., Riesener, M., Schuh, G., 2019. Methodology for the risk and reward evaluation of industrial subscription models, in: Lehnert, Wulfsberg (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 613–622.

[18] Gassmann, O., Csik, M., Frankenberger, K., 2014. The business model navigator: 55 models that will revolutionise your business. Pearson, Harlow, 387 pp.

[19] Wenger, L., 2019. Informationslogistik für Subskriptionsmodelle im Maschinen- und Anlagenbau: Vorstudiengespräch zum Dissertationsvorhaben. FiR an der RWTH Aachen, Aachen.

[20] Stoppel, E., 2016. Nutzungsabhängige Preissysteme auf industriellen Märkten. Springer Fachmedien Wiesbaden, Wiesbaden, 207 pp.

[21] Thiesse, F., Kohler, M., 2008. An Analysis of Usage-Based Pricing Policies for Smart Products. Elec. Markets 18 (3), 232–241.

[22] Decker, C., Paesler, S., 2004. Financing of Pay-on-Production-Models, in: Knorr, A., Lemper, A., Sell, A., Wohlmuth, K. (Eds.), Berichte aus dem Weltwirtschaftlichen Colloquiumder Universität Bremen, pp. 1–15.

[23] Königs, H.-P., 2017. Methoden und Werkzeuge für das Informations-Risikomanagement, in: Königs, H.-P. (Ed.), IT-Risikomanagement mit System. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 247–281.

[24] Probst, G., Gomez, P. (Eds.), 1991. Vernetztes Denken: Ganzheitliches Führen in der Praxis, 2., erweiterte Auflage ed. Gabler Verlag, Wiesbaden, Online-Ressource.

[25] IT-Sicherheitscluster e.V., 2020. Über ISIS12 - InformationsSicherheitsmanagementSystem in 12 Schritten, Regensburg. https://isis12.it-sicherheitscluster.de/ueber-isis12/. Accessed 27 March 2021.

[26] Wadhwa, M., 2020. A Beginners Guide to the STRIDE Security Threat Model. https://www.ockam.io/learn/blog/introduction_to_STRIDE_security_model. Accessed 15 March 2021.

Biography

Prof. Dr.-Ing. Dipl.-Wirt. Ing Günther Schuh (*1958) has held the chair of production systems engineering at RWTH Aachen University since 2002. Furthermore, he is the director of the Institute of Industrial Management (FIR) and a member of the board of directors of the Machine Tool Laboratory WZL at RWTH Aachen and the Frauenhofer Institute for Production Technology (IPT). He is co-founder of the electric vehicle manufacturers Streetscooter and e.GO Mobile.

Prof. Dr.-Ing Volker Stich (*1954) has been 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.

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 intelligent products. As head of the division, he is responsible for multifaceted consulting and research projects.

Lars Kaminski (*1993) started working as a project manager at FIR at RWTH Aachen University in 2019. In his current position within the technical group IT complexity management, he specialized within IT architecture management and the development of transformation strategies concerning information security.

Jacques Engländer (*1992) has been working as a project manager at FIR at RWTH Aachen University since 2018. In his current position as part of the Information Management division, he supports companies in the design of IT strategies and information security management for organizational, technical, and cultural aspects.

Anna Majchrzak (*1990) has been a research student assistant at FIR at RWTH Aachen University since 2020. As a sociologist with an emphasis on the sociology of technology and organization, her primary research interests include organizational development, sociotechnical systems, and the impact of digital transformation in enterprises.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Do We Really Know The Benefit Of Machine Learning In Production Planning And Control? A Systematic Review Of Industry Case Studies.

Kathrin Julia Kramer[1], Alexander Rokoss[1], Matthias Schmidt[1]
[1]*Institute of Product and Process Innovation (PPI)/ Leuphana University, Lueneburg, Germany*

Abstract

The field of machine learning (ML) is of specific interest for production companies as it displays a perspective to handle the increased complexity within their production planning and control (PPC) processes in an economic and ecologic effective as well as efficient way. Several studies investigate applications of ML to different use cases. However, the research field lacks in research on industry case studies. A broad understanding from a practical perspective and in this context, an evaluation from a data mining and business standpoint is key for gaining trust in ML solutions. Therefore, this paper gives a comprehensive overview of evaluation dimensions and outlines the current state of research in ML-PPC by conducting a systematic research overview. First, the present work provides key dimensions of business and data mining objectives as evaluation metric. Business objectives are clustered into economic, ecological and social objectives and data mining objectives are grouped into prediction accuracy, model's explainability, model's runtime, and model's energy use. Secondly, the systematic literature review identifies 45 industry case studies in ML-PPC from 2010-2020. The work shows that the scientific publications only rarely reflect in detail on a wide range of evaluation metrics. Instead, researchers mainly focus on prediction accuracy and seldom investigate the effect of their results to a business context. Positively, some papers reflect on further aspects and can inspire future research. This resulting transparency supports decision makers of companies in their prioritization process when setting up a future ML-roadmap. In addition, the research gaps identified herein invite researchers to join the research field.

Keywords

Machine learning; Production planning; Production control; Process planning; Economic effect; Ecological effect; Practical reference

1. Introduction

In production planning and control (PPC), the focus is on optimizing production logistical objectives (lead-time, inventory, capacity utilization and adherence to delivery dates), which in turn influence the company-wide targets. Henceforward, the fulfilment of PPC tasks occurs in an environment with several interdependencies, which complicates the corresponding decision process [1,2]. In addition, trends such as shorter product life cycles or higher number of product variants further increase the complexity of PPC [2]. The application of machine learning (ML) can support production companies to overcome the increased complexity within their PPC processes and thus positively influences the fulfilment of objectives [3]. In several studies, the application of ML to different use cases was already investigated, which indicate a positive contribution to fulfilling PPC tasks in comparison to previously used methods [3]. Overall, ML

DOI: https://doi.org/10.15488/11296

methods are especially used for tasks with high uncertainty or complexity like demand planning, inventory management, lead time prediction or scheduling [4]. In addition, the automation of simple and repetitive tasks (e.g. preparation of shipping documentation) can be supported with ML methods [5]. Supplementary to this, a growing number of publications in this field of research underpins an increased research interest [4]. Even though several studies exist, it seems that authors rarely apply those solutions to a practical context and seldom provide a comprehensive evaluation of the results [6,7]. This means evaluating developed models and also their impact on existing business processes [8–10]. However, this is key to gaining trust in ML solutions and being able to select an adequate solution in a complex decision-making environment, as is the case with PPC [6,2]. The use of new technologies is not an end in itself, but should be understood in order to be able to use it purposefully [8,6,11]. In conclusion, a clear research gap exists in regards to investigating ML-PPC solutions from a practical perspective and providing transparency on the conducted evaluation. Thus, this paper aims in identifying studies which evaluate ML solutions in a practical context and hence due not base on a simulative environment. With answering the research question '*Do we really know the benefit of ML in PPC?*', the basis for real progress in applying ML-PPC solutions is set as the paper aims for more transparency about practical use cases as well as applied dimensions of evaluation.

The research agenda and the structure of the paper consists of five sections: It starts with forming the research question of the paper as presented in this section. Next, the research subject is described by outlining the current state of research in the field of PPC and ML. The section concludes with a research gap, which this paper addresses and underpins the contribution of this work to the overall ML-PPC research field. In order to answer the research question, section 3 expresses the research methodology. On the one hand, by introducing the method of a systematic literature review and the accompanying search strategy to identify a suitable research sample representative of the research object. On the other hand, by setting an analytical framework for the description of the identified research sample and for the analysis of the evaluation dimensions used by the researchers. Both a business management and a data mining perspective are taken into account. Subsequently, section 4 interprets the results of the applied research methodology. The final part of the research agenda and the last section of the paper summarizes key findings and outlines a future research agenda by presenting research gaps. This is a first step towards a comprehensive evaluation of practical applications of ML methods in PPC and thus enabling real progress in adoption.

2. Current state of research

As the previous section outlines, evaluation is key for understanding the benefits and risks of ML applications. In addition, investigating effects of specific ML solutions in practical contexts can provide important insights about dependencies with other processes. So far, researchers have not systematically reviewed the ways of evaluation across the ML-PPC domain. In this context, research does not consider the linkage between the use of ML methods and business as well as data mining objectives of studied industry case studies in PPC as the following section illustrates.

Usuga Cadavid et al. [7] have analysed 93 ML-PPC papers and assigned those to different use cases of industry 4.0, proposed by Tao et al. [12] and themselves. They give a comprehensive overview and among others outline that the majority of studies uses artificial data and less than half of the papers address an actual implementation of a ML model. A closer investigation of the studies using industry data does not take place. In addition, they do not outline different ways of evaluation [7]. Schmidt et al. [4] focus on ML applications and clustered 94 scientific papers according to PPC tasks of a reference model named Hanoverian Supply Chain Model [1]. However, the authors only show a bibliometric analysis without providing further insights. Other authors have either chosen a more precise aspect in PPC like lead time prediction [13] and scheduling [8,14] or specific methods (e.g. fuzzy methods [15] or deep learning [16]). In addition, reviews exist that have a more holistic view on manufacturing processes [17,16] or are taken within the context of industry 4.0

[18] or cyber-physical production systems [19], without comprehensively displaying ML-PPC applications. Most of those studies do not outline a comprehensive evaluation of practical ML applications. However, Alemão et al., Fahle et al. and Bueno et al. study it partially for their field of study. Alemão et al. highlight within their systematic study about manufacturing scheduling how many and which objectives were reflected in the target function of used reinforcement learning (RL) approaches in the identified studies. They argue that scientific research in developing a solution often focuses on too few aspects to be ultimately practical. The aim of their overview is therefore to lay the foundation for establishing reliable solutions that are suitable for practice [8]. Fahle et al., who cover a wide field of manufacturing processes, centre within their study on industrial use cases using ML or artificial intelligence. Within that scope, they have identified 44 scientific papers. That said, they outline the area of application and the used algorithm, but do not present a structured overview of the respective data mining and business objectives [17]. Bueno et al. focus on PPC in an industry 4.0 context and show as part of their analytical framework that 52 out of 102 analysed studies address performance indicators in relation to industry 4.0. An analysis of the 52 studies displays 13 different performance indicators (e.g. cost, flexibility, productivity, lead time, and customer satisfaction). Most of the identified studies depict the impact of smart PPC on cost, flexibility or reliability. However, ML solutions are only marginally investigated and most of the studied papers do not present the performance in the context of a real manufacturing scenario [18].

In conclusion, an increased research interest in practical solutions can be observed. Nevertheless, none of the reviews so far provides a holistic view on practical ML applications in PPC. Instead, many approaches presented by the scientific community are based on simulation data and do not translate to a business environment [7,11]. A comprehensive presentation of industry use cases using ML applications within PPC accelerates future research and eases the adoption by manufacturing companies. By outlining different evaluation instruments for ML approaches in PPC, decision makers of companies are able to prioritize ML use cases in their company more suitable. Thus, this paper conducts a systematic research overview with attention on scientific papers, which record ML approaches within real business processes. There is a particular emphasis on identifying the progress of the scientific community in reflecting on various objective dimensions, as evaluation is fundamental to making actual progress in data mining [10] as well as to receive effective and efficient PPC processes. Especially in PPC, as several conflicting goals exist, a global consideration of goals is important to avoid suboptimal optimizations [1].

3. Research methodology and analytical framework

To meet the objectives of this study, a systematic literature review is conducted to draw insights from scientific literature about existing ML solutions in PPC. As an immense amount of publications exist in the ML-PPC research domain, systematic literature reviews can give a comprehensive overview of the state of research. Systematic reviews provide insights from existing scientific papers and support the identification of research gaps and common patterns. Aims are to enhance the knowledge base on a concrete research topic and enlighten practitioners as well as policymakers [20]. The second part of the research methodology foresees an analysis of the final publications' sample. This analytical framework consists of three questions that contribute to the overall research question of this work.

The literature review bases on the method suggested by Tranfield et al. [20], which was already successfully employed by other authors of the domain [4,7]. As stated in section 2, there is no analysis of the scientific ML-PPC literature in a practical context. The research was performed between 17/12/2020 and 26/01/2021 in the established databases Scopus and Web of Science [21]. Figure 1 summarizes the seven steps of the systematic literature review.

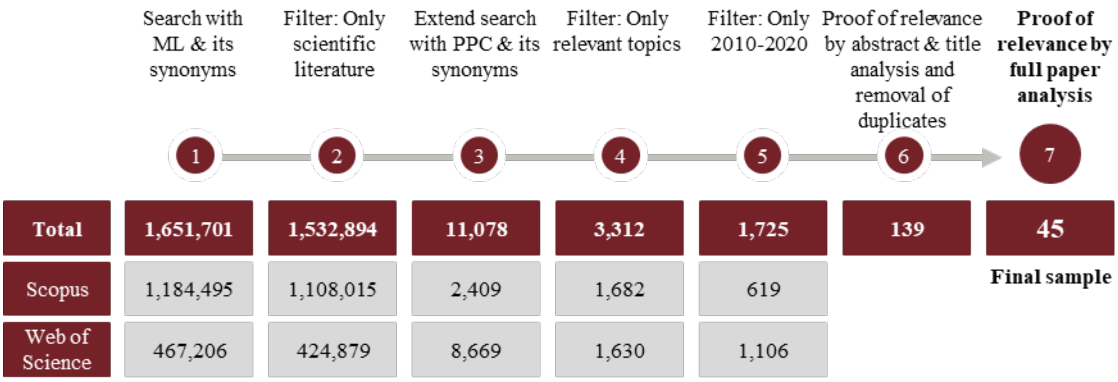


Figure 1: Search procedure of the scientific literature review

The process begins with searching titles, keywords and abstracts of publications within the research field 'machine learning' as well as corresponding synonyms. In industrial practice 'artificial intelligence', 'data analytics', 'data mining', 'deep learning' and 'neural networks' often function as synonyms to ML and are therefore part of the search string [9,10]. The second step foresees a limitation to only scientific literature to ensure high-ranked publications. Henceforward, the search is limited to journal articles and conference papers. In the third step, PPC related terms are added with an AND-function to the search string. As Schmidt et al. point out common PPC terms are 'production planning', 'production control', 'production scheduling', and 'process planning' [4]. In the next step, only publications of subject areas, which relate to PPC in production environments, are selected. This is achieved by setting corresponding filters of subject areas (e.g. 'engineering') in the databases. In order to facilitate an in-depth analysis – in the fifth step of the search procedure – the sample is limited to scientific papers from 2010-2020. The sixth step foresees to check if abstracts and titles of the remaining publications are mentioning a ML-PPC topic within a production environment. Otherwise, they are excluded. Full papers are necessary for further analysis; consequently, only papers that are available online are part of the final sample. After removing duplicates between the results from Scopus and Web of Science, a sample of 139 scientific publications remains. It should be noted at this point that the assumptions made in conducting the literature search certainly do not lead to a universal identification of all relevant papers. Nevertheless, it can be stated that a representative number of papers were identified and thus a sufficient sample for answering the research question is hereby available. In the final step, the sample is further prepared for performing a detailed analytical framework. Hence, it is necessary to eliminate papers, which do not cover a case study based on real data. Authors who show a simulation-based case study or who do not present a case study cannot contribute to the set research question. It is noticeable that 44.6% of papers show a sole use of simulation data and 23.0% do not examine a case study. In total, the final sample consists of 45 publications. Following, to gain knowledge and insight from the remaining papers, an exploration involves a full paper analysis. This analytical framework consists of three questions that contribute to the overall research question of '*Do we really know the benefit of ML in PPC?*'.

The first question is *'How are the identified studies characterized in regards to their used learning type of ML, the year of publication and the applied use case?'*. By answering those questions, a frame of the final sample is set. First, to make the vast number of ML algorithms more comprehensible, they can be clustered according to their learning type. Common learning types are supervised learning (SL), unsupervised learning (UL) and RL: SL methods use input factors with assigned output as a learning basis. The model learns to recognize and generalize the relationship between input and output data and uses this to predict outputs for unknown examples. In contrast, methods of unsupervised learning are characterised by unknown output data and serve to categorize or segment data. Methods of reinforcement learning learn to take a decision based on feedback loops [9]. Second, by recording the year of publication, it is possible to explore whether the number of publications by authors engaged in real-data case studies has changed over time. Third, applied use cases in PPC are classified. Accordingly, it shows whether some use cases have been dealt more than others have. The analysis bases on the main sub tasks in PPC as the PPC process normally constitutes of several sub tasks in order to handle the planning complexity as well as the different planning horizons (short-, mid- and long-term) [19,2]. In this paper, similar to Schmidt et al. [4], the sub tasks base on the Hanoverian Supply Chain Model and consist of eleven main tasks for fulfilling an efficient and effective planning, control and monitoring of production orders [1].

Afterwards, a structured overview of the reflected data mining and business objectives in practical ML-PPC case studies is presented for providing insights about the final sample of publications. This leads to the next two questions of the analytical framework: *'How many studies have addressed data mining and business objectives? Which data mining and business dimensions were displayed?*. To answer the set questions, clusters for data mining and business objectives are formed, as shown in Figure 2. To begin with, scientists evaluate developed ML models by using different data mining objectives. Common objectives, which form the basis for the analytical framework of this paper, are prediction accuracy, model's explainability, model's runtime, and use of energy [22,9,10]. Prediction accuracy of a model can be measured using various performance measures (e.g. receiver operating characteristic curve (ROC-curve), mean-squared error or root mean-squared error (RMSE)), depending on the problem at hand [10]. Further explanations can be found in [10], [23] and [24]. Evaluating explainability of a model means to reflect on traceability and understandability of a model's result as well as its solution path [9,10]. Methods generally differ between interpretable white-box models (e.g. decision tree) and ex-post explanations of black-box models (e.g. explanations of attributes) as further explained in [25] and [26]. The model's runtime can be differed between runtime during training and during execution of a model. In this course, runtime highly depends on the used hardware and the model's complexity [27]. As the product of time and power is energy, model's energy consumption also belongs to the presented evaluation dimensions [22,9]. A broad overview on different energy-related evaluation methods is provided in [22].

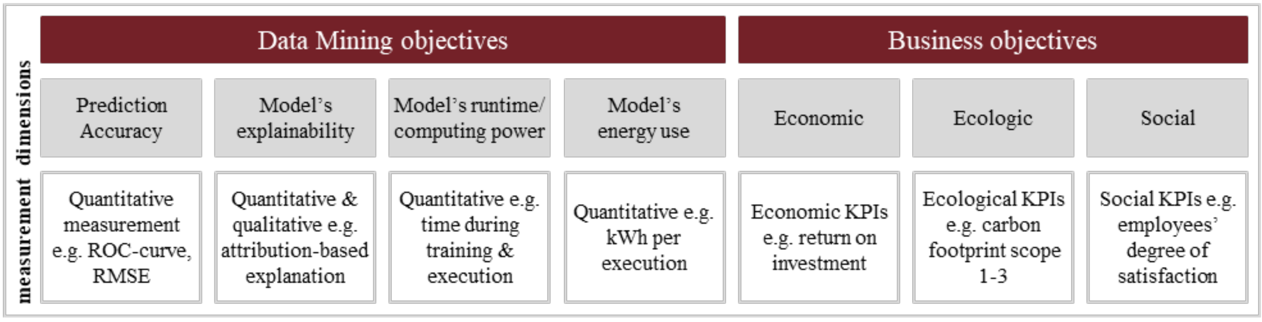


Figure 2: Dimensions of data mining and business objectives, after [28,29,22,9,10]

Next to data mining objectives, a further linkage to an actual business context is as important. In this paper, as shown in Figure 2, numerous business objectives (e.g. return maximization, minimizing carbon emissions) are clustered according to their economic, ecologic and social effect. Business objectives are directed either to one specific cluster or to several ones. In addition, companies usually pursue different business objectives at the same time, which can be neutral, complementary or contradictory to each other [29]. Economic factors usually make up the core of companies' objectives and concentrate on return on investments and its influencing components. Thus, in manufacturing a better process performance is targeted [28]. Nevertheless, as several scientists stress, ecological and social factors are as important as economic factors for attaining a sustainable business [29,7]. Within a ML-PPC context, Usaga Cadavid et al. emphasize that human collaboration and environmental traits should be taken into account when developing a ML model [7]. The ecological dimension contains objectives that seek to reduce the environmental impact. In manufacturing, this is achieved through an improved viable use of resources (e.g. energy, material) [28]. In the light of global ambitions to reduce carbon emissions (e.g. European Green Deal), the ecological dimension might gain additional interest in the near future [30]. Finally yet importantly, the social dimension includes effects on humans (e.g. employees, suppliers) like ethical consequences [31]. Hence, the aim is to reflect on improved work conditions for employees in manufacturing [28]. In a ML-PPC system, for example, the impact on labours' working conditions can be judged [7]. Finally, it should be noted that papers are only counted if they clearly address this linkage in their application section in either a quantitative or a detailed argumentative manner.

Overall, the analytical framework enables a comprehensive investigation of the final sample of the conducted literature review. This paper presents practical use cases in PPC in a transparent way and particularly illustrates the linkage with data mining as well as business management goals. This leads to a comprehensive tool kit for evaluation. In addition, thanks to the research methodology research gaps and areas of focus within the ML-PPC research domain are highlighted.

4. Results

The conducted systematic literature review leads to 45 scientific publications covering real-data-based case studies from ML-PPC research fields from 2010-2020. The fact that many papers were excluded during the review process is in line with the thesis of Usuga Cadavid et al. [7] and Wuest et al. [11], who outline that publications often do not use data from industry use cases. Figure 3 illustrates the results of the first part of the analytical framework, introduced in chapter 3.

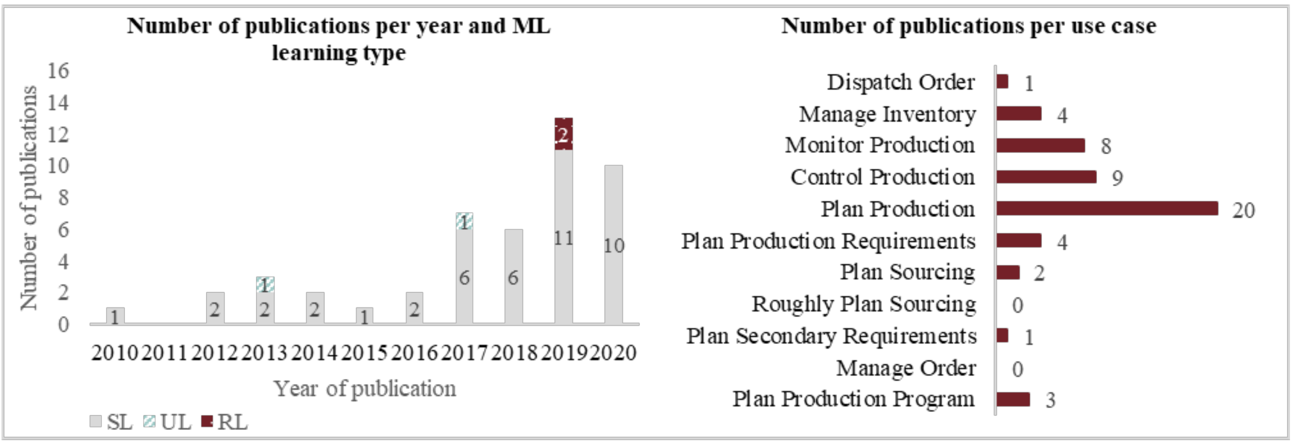


Figure 3: Publications of final sample clustered by ML learning type, publication year and PPC use case

The left side of Figure 3 shows the number of publications per year and ML learning type. From 2010 to 2016, less than three publications apply ML to a PPC use case in an industry context. From 2017 onwards, the number of publications steadily increased with a peak in 2019. Looking at the used ML learning type, mostly supervised learning approaches were identified. Most cases did not combine a learning type with another learning type. The right side of Figure 3 illustrates the number of publications per use case. It clearly displays that the scientific research community mainly applies ML models to 'Plan Production', 'Control Production' and 'Monitor Production'. Within this, to state the three most common sub tasks, 16 publications deal with lead time prediction and its influencing components (for example [32–34]), 8 publications cover production control (for example [35–37]) and 4 publications address sequencing (for example [38]).

Next, Figure 4 and Figure 5 visualise the results of the second part of the analytical framework, that form the key contribution of this paper. The left side of Figure 4 shows the covered evaluation aspects in relation to a ML model itself. It clearly states that most studies (51.1%) purely focus on prediction accuracy. In total, 95.6% of the sample evaluated on prediction accuracy, 33.3% on model's explainability, 22.2% on model's runtime, and none on energy use. The right side of Figure 4 presents the counted publications in regards to reflected dimensions of business objectives. 80.0% of the sample do not evaluate on business objectives.

The remainder reflects on economic (20.0%) and ecologic (2.2%) aspects. Social aspects were not addressed in detail.

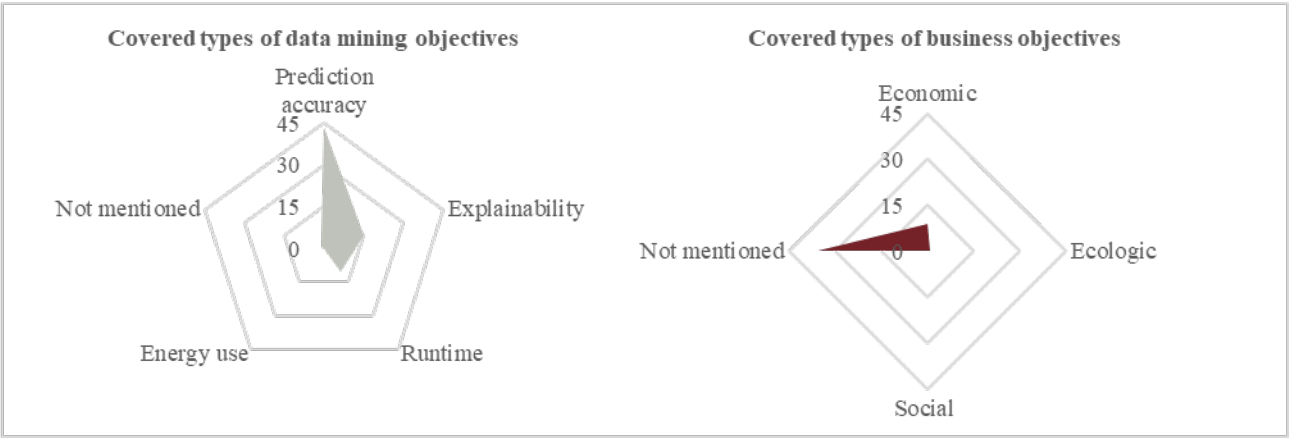


Figure 4: Publications of final sample grouped by types of data mining and business objectives

Figure 5 illustrates how comprehensive the authors evaluate their industry studies. It expresses that most often one or two data mining objectives and zero to one business dimension were reflected. None of the identified papers comprehensively investigates on economic, social and ecologic objectives and simultaneously broadly studies different data mining objectives. In regards to data mining goals, almost all papers (97.8%) evaluated on one data mining objective and 44.4% of the sample reflected on more than one data mining objective. The three dimensions of business objectives are underrepresented as only 9 out of 45 papers address business objectives. A paper of Moerzinger et al. is the only paper that covers two dimensions of business objectives and two data mining goals [39].

Business objective / DM objective	No dimension covered	One dimension covered	Two dimensions covered	Three dimensions covered
No type covered	0	1	0	0
One type covered	19	5	0	0
Two types covered	13	2	1	0
More then three types covered	4	0	0	0

Figure 5: Relationship between data mining types and dimensions of business objectives of the final sample

In conclusion, this section framed the identified studies and outlined a structured overview of reflected data mining and business objectives in industry ML-PPC use cases. The next section considers the results for presenting research gaps and similarities.

5. Conclusion and future research agenda

This work has identified different dimensions of evaluation and has outlined the current state of research in ML-PPC industry case studies. However, two-thirds of ML-PPC studies, published between 2010 and 2020, do not cover real industry cases. A reason could be that it might be difficult for researchers to get access to production data as companies often classify them as confidential. Nevertheless, researchers should aim for applying their developed models to real use cases, as a reality gap might exist between simulation models and real environments. Henceforward, industry use cases form the basis for concluding on potential benefits of ML solutions and for answering the set research question. In total, 45 studies of industry case studies in ML-PPC were found. Within this scope, the motivation of the paper was twofold.

First, the conducted systematic review presents that mainly the PPC tasks ‘Plan Production’, ‘Control Production’, and ‘Monitor Production’ were reflected in real data case studies. This is in line with the findings of Schmidt et al. who investigate this for simulation-based as well as real data-based case studies from 2009-2019 [4]. Still, future research should also aim to address other PPC tasks (e.g. ‘Plan Production Program’ or ‘Manage Inventory). In addition, publications about sequencing are rare herein. It shows that this often addressed research field still lacks in industry application, probably due to a wide spread of RL solutions. In order to apply RL models to real industry cases, a sufficient digital twin is necessary, as the cost for training in a real environment would not be sufficient [10,11]. This might explain the rare findings of RL cases within the final sample. SL models are mostly presented, as several regression and classification tasks exist within PPC [7,11]. In regards to development over time, it can be anticipated, that in line with the generally increasing research interest in ML-PPC [3,4], real industry applications also gain in importance.

Secondly, a structured overview of reflected data mining and business objectives in practical ML-PPC case studies was prepared. The work shows that most researchers do not comprehensively evaluate their ML models in order to conclude on benefits of ML sufficiently. The systematic review displays that ML models were mostly evaluated purely in regards to their prediction accuracy. However, the cost in the sense of time and computation power was regularly not assessed. With an increasing number of deployed ML models within different industries, the energy use should not be disregarded. In addition, time of training gains in a more frequently changing production environment of importance and time of execution is of significance especially for real-time tasks (e.g. in production control) [11]. Thus, clear research gaps exist in these matters. Next to the evaluation of models itself, the models’ effect on business objectives is even more underrepresented in the final sample. In total, 80% of the papers do not reflect on business objectives. In PPC, this is of risk, as several contradictory objectives exist: On the one hand, between the economic, social and ecological dimensions and on the other hand also within dimensions. As from the current nature of PPC, research is economic driven [1]. This might explain why those business objectives are more often presented in the identified studies then social and ecological objectives. Nevertheless, future research should consider social, ecological and economic implications for accomplishing a viable business model [29]. Further, as ML-PPC case studies generally focus on specific sub tasks in PPC, it is of importance to reflect on consequences towards other tasks [1,11]. Otherwise, a risk occurs that only sub optimal solutions and not a global optimum is found. In conclusion of this paper, an answer to the set research question – *if we really know the benefit of ML in PPC?* – is given as follows: Currently, due to a lack of a broad evaluation, researchers and practitioners cannot yet clearly state the benefit of ML in PPC. Thus, more industry case studies across different ML learning types and use cases are needed in future research and researchers should evaluate their results broadly. Consequently, the findings can be further developed towards a conceptual framework that enables a broad evaluation metric. This can be supplemented by the state of research from different research fields. Further, a summary of different key performance indicators as well as methods of measurement can enrich this overview.

The work delivered transparency on the current state of research in real ML-PPC use cases. For the future, it can be concluded which dimensions should be considered additionally, when developing ML-PPC solutions. An awareness of different dimensions of evaluation was created, that supports decision makers of companies in their prioritization process when setting up a future ML-roadmap. As previously stated, evaluation is important to making real progress [10], thus researches should join the research field and focus on the research gaps identified herein.

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] Schmidt, M., Schaefers, 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.

[2] Wiendahl, H.-P., 2014. Betriebsorganisation für Ingenieure, 8th ed. Carl Hanser, München.

[3] Bullers, W. I., Nof, S. Y., Whinston, A.B., 1980. Artificial Intelligence in Manufacturing Planning and Control. AIIE Transactions 12 (4), 351–363.

[4] Schmidt, M., Maier, J. T. & Grothkopp, M., 2020. Eine bibliometrische Analyse: Produktionsplanung und -steuerung und maschinelles Lernen. wt Werkstattstechnik online 110 (4).

[5] Knoll D., Neumeier D., Prüglmeier M., Reinhart G., 2019. An automated packaging planning approach using machine learning. Procedia CIRP 81, 576–581.

[6] 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).

[7] 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. J Intell Manuf 31 (6), 1531–1558.

[8] Alemão, D., Rocha, A.D., Barata, J., 2021. Smart Manufacturing Scheduling Approaches—Systematic Review and Future Directions. Applied Sciences 11 (5), 2186.

[9] Jordan, M.I., Mitchell, T.M., 2015. Machine learning: Trends, perspectives, and prospects. Science 349 (6245), 255–260.

[10] Witten, I.H., Eibe, F., Hall, M.A., 2011. Data Mining: Practical Machine Learning Tools and Techniques, 3rd ed. Elsevier.

[11] 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.

[12] Tao, F., Qi, Q., Liu, A., Kusiak, A., 2018. Data-driven smart manufacturing. Journal of Manufacturing Systems 48, 157–169.

[13] Burggraef, P., Wagner, J., Koke, B., Steinberg, F., 2020. Approaches for the prediction of lead times in an engineer to order environment – a systematic review. IEEE Access 8, 142434–142445.

[14] Lin, L., Gen, M., 2018. Hybrid evolutionary optimisation with learning for production scheduling: state-of-the-art survey on algorithms and applications. International Journal of Production Research 56 (1/2), 193–223.

[15] Wong, B.K., Lai, V.S., 2011. A survey of the application of fuzzy set theory in production and operations management: 1998–2009. International Journal of Production Economics 129, 157–168.

[16] 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, 144–156.

[17] Fahle, S., Prinz, C., Kuhlenkötter, B., 2020. Systematic review on machine learning (ML) methods for manufacturing processes – Identifying artificial intelligence (AI) methods for field application. Procedia CIRP 93, 413–418.

[18] Bueno, A., Godinho Filho, M., Frank, A.G., 2020. Smart production planning and control in the Industry 4.0 context: A systematic literature review. Computers & Industrial Engineering 149, 106774.

[19] Gallina, V., Lingitz, L., Karner, M. (Eds.), 2019. A New Perspective of the Cyber-Physical Production Planning System.

[20] 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, 207–222.

[21] 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.

[22] García-Martín, E., Rodrigues, C.F., Riley, G., Grahn, H., 2019. Estimation of energy consumption in machine learning. Journal of Parallel and Distributed Computing 134, 75–88.

[23] Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer New York, New York, NY.

[24] Armstrong, J., Collopy, F., 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting 8 (1), 69–80.

[25] Zhou, J., Gandomi, A.H., Chen, F., Holzinger, A., 2021. Evaluating the Quality of Machine Learning Explanations: A Survey on Methods and Metrics. Electronics 10 (5), 593.

[26] Goldman, C.V., Baltaxe, M., Chakraborty, D., Arinez, J., 2021. Explaining Learning Models in Manufacturing Processes. Procedia Computer Science 180, 259–268.

[27] Kaxiras, S., Martonosi, M., 2008. Computer Architecture Techniques for Power-Efficiency. Synthesis Lectures on Computer Architecture 3 (1), 1–207.

[28] Braccini, A., Margherita, E., 2019. Exploring Organizational Sustainability of Industry 4.0 under the Triple Bottom Line: The Case of a Manufacturing Company. Sustainability 11 (1), 36.

[29] Dyllick, T., Hockerts, K., 2002. Beyond the business case for corporate sustainability. Bus. Strat. Env. 11 (2), 130–141.

[30] Commission to the European Parliament, European Council, Council, European Economic and Social Committee, Committee of the Regions, 2019. Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - The European Green Deal: COM/2019/640 final.

[31] Russell, S.J., Norvig, P., 2010. Artificial Intelligence: A Modern Approach, 3rd ed. Pearson Education, Upper Saddle River, New Jersey.

[32] Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., Monostori, L., 2018. Lead time prediction in a flow-shop environment with analytical and machine learning approaches. IFAC-PapersOnLine 51 (11), 1029–1034.

[33] Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., Monostori, L., 2018. Lead time prediction using machine learning algorithms: A case study by a semiconductor manufacturer. Procedia CIRP 72, 1051–1056.

[34] Wang, C., Jiang, P., 2019. Deep neural networks based order completion time prediction by using real-time job shop RFID data. J Intell Manuf 30 (3), 1303–1318.

[35] Ma, S., Zhang, Y., Lv, J., Ge, Y., Yang, H., Li, L., 2020. Big data driven predictive production planning for energy-intensive manufacturing industries. Energy 211, 118320.

[36] 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.

[37] Wang, C., Jiang, P., 2018. Manifold learning based rescheduling decision mechanism for recessive disturbances in RFID-driven job shops. J Intell Manuf 29 (7), 1485–1500.

[38] Song, J., Song, X., Wang, X., 2019 - 2019. A New Scheduling Model for Tire Production and Transportation Among Distributed Factories, in: 2019 15th International Conference on Computational Intelligence and Security (CIS). 2019 15th International Conference on Computational Intelligence and Security (CIS), Macao, Macao. 13.12.2019 - 16.12.2019. IEEE, pp. 50–53.

[39] Mörzinger, B., Loschan, C., Kloibhofer, F., Bleicher, F., 2019. A modular, holistic optimization approach for industrial appliances. Procedia CIRP 79, 551–556.

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.

Alexander Rokoss (*1993) 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 2018. His 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 Lueneburg. In addition, he became the head of the PPI in 2019.

2nd Conference on Production Systems and Logistics

Enabling The Smart Factory – A Digital Platform Concept For Standardized Data Integration

Julia Donnelly[1], Andreas John[2], Jonas Mirlach[3], Kilian Osberghaus[3], Silvia Rother[4], Christian Schmidt[2], Hannes Voucko-Glockner[3], Simon Wenninger[5]

[1]FIM Research Center, University of Augsburg, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Germany
[2]soffico GmbH, Augsburg, Germany
[3]FIM Research Center, University of Augsburg, Germany
[4]University of Augsburg, Germany
[5]FIM Research Center, University of Applied Sciences Augsburg, Project Group Business & Information Systems Engineering of the Fraunhofer FIT, Germany

Abstract

The digital transformation of industrial manufacturing companies is still seen as a challenge on the path to the smart factory and efficient and transparent production. Companies have already been dealing with concepts such as Industry 4.0 and the associated tasks of digitization and process automation for years. More recently, data-driven methods of machine learning and artificial intelligence, have increased the demands on data integration in industrial applications and the need for seamless and automated communication within information systems. This facilitates the creation of entirely new business models or the adaption of new approaches such as predictive analytics (e.g. for predictive maintenance) or data mining methods (e.g. for anomaly detection) to increase productivity. A multitude of partly proprietary standards for communication between machines and systems, and interface definitions complicates the integration of systems and data. To minimize the challenges of system and data integration, we have developed a digital platform concept as a solution to standardized data integration issues. In a first step, through literature research and expert interviews, we identified current industrial trends and key relevant standards and processes. In a second step, we developed a concept for a digital platform - a Service Hub - through which the identified standards and integration processes can be marketed. The Service Hub supports companies in their digital transformation and offers providers of various integration solutions an opportunity for individualized marketing of frequently used (integration) processes.

Keywords

Smart Factory; Data Integration; Digital Transformation; Industry 4.0; Digital Platform

1. Introduction and method

New digital technologies and Industry 4.0 are having a significant impact on companies' strategies and business models, as well as on their processes and routines. On the way to efficient and transparent production in the smart factory, companies must perform a digital transformation to remain competitive in international markets. In the context of Industry 4.0, data and the connectivity of machines and systems across all levels of automation play a central role [1]. The gathered volumes of data open up entirely new opportunities for business models and potential for efficiency gains and flexibility [2]. For instance, applications of machine learning and artificial intelligence methods are becoming increasingly important in areas such as process mining or predictive maintenance [3]. These applications require data integration and workflows for the automated and flawless communication of various information systems. However, despite existing technology and the use of middleware software, manufacturing companies still face major problems in integrating data-driven applications for optimisation and automation tasks [2]. A multitude of partly proprietary communication standards and interface definitions as well as non-convenient provision of integration solutions are considered to be key challenges in data integration [2]. Platform approaches to this problem already exist, but they are often specifically tailored to the products and services of the respective provider and use proprietary interfaces and communication standards instead of open standards [4–6]. However, it affects the interaction with external systems and thus counteracts the generative approach of the platform and leads to lock-in effects [5]. To address this issue, we formulate our guiding research question (RQ) as follows:

How can software vendors improve data integration offerings using a digital platform to facilitate data integration and enable a smart factory for manufacturing companies?

We answer the RQ by developing and designing a conceptual model of a digital platform which supports companies in their digital transformation by offering (data) integration solutions via individualized marketing of frequently used (integration) processes. In doing so, we rely on existing solutions and classify our solution and our contribution to practice and the theoretical body of literature as "exaptation", which is characterized by an extension of known solutions to new problems [7]. The aim of the work is to develop a concept that helps companies to circumvent the problem of data integration. This concept can then be a basis for products and services of software vendors offering data integration solutions. Therefore, research in this paper has been organized by an adapted design science in information systems research framework combining behavioural science and design science paradigms [8]. In a first step we identified current industrial trends and relevant standards and processes for typical data integration tasks as a basis for our work by conducting a semi-structured literature review in the databases Science Direct, AIS eLibrary, Springer, and Research Gate with the keywords “IoT”, “Industry 4.0” “platform”, and “data integration”. We then enrich the findings from literature with additional expert interviews and compare them to practical issues to address the needs of manufacturing companies. We therefore reviewed 10 experts who were asked to give their assessment of current trends and the status quo of data integration in (their) industry. These experts work either as consultants or in manufacturing companies and have prior knowledge in the subject matter covered. Using both perspectives allows us to identify and highlight similarities and differences from research a practice. These compiled findings are then used to develop our conceptual model. We structure our paper with the topic of drivers, trends, and standards for data integration in chapter 2, before also describing existing approaches to data integration and providing an overview of digital platforms, with the topics of cloud computing and payment models being highly relevant. Subsequently in chapter 3, we develop a concept for a digital platform - a Service Hub - through which the identified standards and integration processes can be commercialized present our developed Service Hub and describe its cloud-based usage and billing models. Both perspectives from the first two steps contributed to the design phase as well as in the evaluation and refinement of the Service Hub. In chapter 4, we discuss its usage in business, limitations and prospects for further research before we conclude in chapter 5.

2. Theoretical background

2.1 Drivers, trends, and standards for data integration in manufacturing companies

Analyzing literature on current drivers and trends in manufacturing shows the relevance of a seamless data integration. Current trends in manufacturing are primarily shaped by the digital transformation which has been driven by the Internet of Things, Internet of Services, and Big Data for more than a decade. It enables organizations to create novel business models and thereby achieve competitive advantages and outperform within its own industry [9,10]. Nowadays, companies can in particular establish competitive advantages by exploiting technological enhancements in terms of managing and orchestrating data and information. This includes the integration of cyber physical systems, artificial intelligence, machine learning or Distributed Ledger Technologies, such as Blockchain applications within internal processes [10]. In fact, companies are evolving from product providers to service providers through these data-driven technologies. Data-driven analyses or data mining methods, for instance, predictive analytics, especially predictive maintenance and predictive sales require the efficient integration of relevant data to provide real-time forecasts of future developments and risks [11]. In addition, data mining methods enable the identification of valuable new patterns from data. e.g. in customer relationship management (CRM) systems or enterprise resource planning (ERP) systems, leading to increased efficiency and flexibility of a company.

Data or information integration is understood as the consolidation of redundancy-free data and information from different application areas and systems. This requires extracting and evaluating the same data from different systems and different data from the same systems to provide the data to the correct users/systems at the right time [12]. As systems are often incompatible, interfaces are the key for data integration. A middleware is a software abstraction layer that provides those interfaces to help manage the complexity and heterogeneity between different systems [13]. As a result companies can benefit from a higher data integrity, referred to as the correctness and completeness of data, and can thus optimize information flows, process integration, and decision-making in an enterprise [14].

Data integration is considered an ongoing task of information system (IS) management, such as CRM or ERP systems, and the administration of database systems [14]. For information exchange the Extensible Markup Language (XML) and Java Script Object Nation (JSON) have established themselves as a text-based standard structuring language describing the structure of a data set [15,14]. Both may include additional records, enabling the creation of deep nested structures, and allow the exchange and processing of data between software and hardware applications. XML, JSON, and other standard structuring languages therefore form the basis for many software standards.

Table 1: Selection of frequently used data integration standards in the industry

Abbreviation	Notation	Description
-	MTConnect	An XML-based standard which retrieves process information from manufacturing machines via Transmission Control Protocol connections.
HTTP	Hypertext Transfer Protocol	An internet protocol that enables data transfer and communication between internet-connected applications using Representational State Transfer architectures.
OPCUA	Open Platform Communications Unified Architecture	A machine-to-machine communication technology standard based on which multiple devices, sensors and machines can be linked via a single communication thread.
MQTT	Message Queue Telemetry Transport	An open protocol to enable the exchange of data by transmitting data in the form of messages to multiple clients.
AMQP	Advanced Message Queuing Protocol	An open standard for business communication, which enables the transfer of information between companies and applications via a network protocol.

Table 1 lists a selection of frequently used standards for data integration in industry and business based on literature. While every private web user worldwide uses Hypertext Transfer Protocol (HTTP) links when accessing web browsers, OPC UA, MQTT and AMQP represent standards for industrial and business-relevant applications. These standards should be interoperable with the digital platform concept.

To enrich our above findings from literature we also conducted expert interviews. 10 experts from industry were questioned on the current status quo, development and actions taken within their companies with regard to data integration as well as their outlook what additional action should be taken, whether this should be done in the short or long term and where barriers arise in the short and long term. The analysis of the current state revealed that only half of the respondents are intensively dealing with data integration and only 30% have integrated a general standard such as OPC UA. Alternative predominantly proprietary solutions are still widely used. Nevertheless, all experts recognize the need for increasing data integration and its associated potential. Potential is primarily expected in newly enabled interdisciplinary information flows, which will allow for deeper insights and knowledge to be gained and, at the same time, a better understanding of how this knowledge can be protected and applied. In addition, an expected higher data quality offers advantages in terms of efficiency and cost by, for example, automatically identifying and correcting erroneous values. Among the barriers preventing more advanced data integration, both organizational and technical capabilities are considered. From an organizational perspective, the lack of awareness for data integration and its benefits for all stakeholders of a company is a key barrier. In addition, the management levels have often not yet been sensitized to the issue. At the same time, prevailing silo structures must be broken down to enable holistic cross-departmental data integration which requires management support. Moreover, companies often lack technical know-how and transparent documentation of existing processes, systems, and data flows. From a technical perspective, reasons for the lack of overarching data integration standards are primarily seen in historically and often isolated implemented systems with proprietary solutions. However, the simultaneous integration of internal and external data sources while taking data protection principles into account also often proves difficult. In contrast, financial hurdles are currently the smallest barrier to efficient data integration. Most experts agree that the topic of data integration must be addressed in the short term, in a period of less than six years, in order not to be displaced by competitors within the industry. Here, the technical capabilities mentioned above represent the greatest challenge. In the long term, data protection requirements, know-how protection, data security and the allocation of responsibility in the event of data loss are considered to be the main challenges, while at the same time complying with regulatory requirements.

Whereas the literature analysis gives an overview of a variety of possibilities for data and information integration, the evaluation of the survey reflects the current state in the industry. Theory and practical application are currently still far apart. The literature research revealed that functional standards such as MQTT already exist, but that they are more of a wish than a reality in application. Despite this, there is an evident willingness to change. Standardized solutions for connecting and integrating machines can support manufacturing companies in their digital transformation to Industry 4.0 standards and reduce challenges. Furthermore, data integration is necessary to establish new business models and thus to remain competitive, emphasizing the relevance of the Service Hub.

According to literature, several approaches for simplifying data integration already exist, however these approaches seem rather technically focused and mostly concern the standardization of protocols [5]. For example, Sanderson et al. [12] introduce an architectural data integration approach for production enterprises using data distribution services whereas Schuh et al. [16] present the "Internet of Production" as an enabler of cross-domain collaboration for production, development and usage data. With the success of digital platforms, e.g. as app stores in the consumer sector, digital platforms are also increasingly being used in industry and promise future potential for various applications [1].

2.2 Digital platforms

Digital platforms are an emerging and ubiquitous phenomenon in the private sphere as well as across industries [17]. Digital platforms change the way digital products and services are consumed, delivered, and mediated. Established companies are being forced to innovate and transform their current business model to remain competitive. In contrast to the common value creation within companies or linear supply chains, digital platforms enable the co-creation of value in an ecosystem of independent actors and stakeholders [17]. Constantly evolving information technology, such as cloud computing and analytic solutions for Big Data, provide the technological foundation for digital platforms and enable entirely new business models, distribution opportunities as well as payment and tariff models [18,19].

The ecosystem of digital platforms typically consists of a platform provider, one or more value-adding mechanisms, and complementors. The platform provider implements governance mechanisms to enable value-creating mechanisms on a digital platform between the platform provider and an ecosystem of autonomous complementors and customers [17,20]. Value-creating mechanisms are divided into two overarching mechanisms and are central for the success of digital platforms. First, there is the value-creating mechanism of transaction facilitation describing the digital interaction between complementors and consumers by, for example, directly matching supply to demand and proposing possible transactions or offers to consumers. The second value-creating mechanism is the mechanism of enabling complementors to jointly create synergistic solutions. Complementors can be providers of different products or services [21,22].

From a technical point of view, a digital platform is a software-based platform that, in addition to basic functionality, provides the option of adding modular services, replacing them, and maintaining a stable core. Each module extends the functionality of the platform and may be implemented by both external parties (complementors and platform provider). Furthermore, the design of digital platforms creating a two-sided or multi-sided market, combined with highly customizable digital technologies, enables high levels of scalability by leveraging network effects [23]. This offers digital platforms the opportunity to easily acquire new customers and to grow with existing customers.

In today's digital society, cloud computing is indispensable. Cloud services gradually replace on-premise systems for data processing in companies and also represent a key technology for digital platforms [23]. Cloud computing enables on-demand computing services to be provided with high reliability, availability, and scalability in a distributed environment [24]. A variety of different service models already exist in cloud computing, which allow the customer to purchase individualized and highly customized infrastructure.

These individualized and highly customized infrastructures can be priced very differently which is reflected in the development of various new and innovative payment models with the ongoing digitalization leading to digital hybrid value creation processes and a shift from product-only providers to service providers [25]. With payment models being a relevant component of business models, we conducted a literature research on different payment models and compared them to the results of the conducted expert interviews, allowing us to enhance the design of the Service Hub. Considering digital transformation and customer needs, platform providers are primarily focusing on pay-per-use models, subscription models, and one-time purchases. In a pay-per-use model, the user pays according to his individual scope of use resulting in individual prices for each user [26]. There are no acquisition costs, capital commitment or software maintenance costs [27]. In a subscription model a fixed price is paid on a regular basis and does not depend on the actual scope of use of the product [28]. In the case of a one-time purchase, ownership of the product is transferred to the buyer who can use the product for its entire lifetime.

Following the literature review on the three predominant payment models, we questioned the experts about their preference to take the needs of potential platform customers into account for our newly developed concept of a digital platform. We distinguished between payment models for a basic version of a middleware and additional individual modules for specific data integration processes. 75% of the experts consider Pay-per-use and subscription models as preferred models for the basic version, whereas pay-per-use was favoured by the majority for additional modules. In addition to the selection of payment models the experts' choice is strongly dependent on the scope of services, including support and updates, as well as the scope of use. If support and updates are included in the basic version, acceptance of the subscription model increases compared with pay-per-use. If the product is not guaranteed to be up to date, experts recommend that the pay-per-use model be degressive, resulting in a choice of this payment model. Consequently, the scope of services and the scope of use must be defined in advance for the design of various payment models, which differ between the basic version and additional modules.

3. A digital platform concept for standardized data integration

With the RQ initially defined, we developed a conceptual model of a digital platform - the Service Hub - based on the presented results from the literature research and the expert interviews. It may support companies in their digital transformation by offering (data) integration solutions via individualized marketing of frequently used (integration) processes. The main idea of our solution is based on that of established app stores. In these app stores, customers buy applications, which can be implemented and used on the customers' (mobile) device. In the business-to-business area it works the same way but individualized and highly specified data integration solutions are required. The design of the digital platform concept from a business model perspective is conceptually illustrated in Figure 1. It incorporates the three central elements platform provider, complementor, and value creating mechanisms. In addition, the model depicts all relevant connections, dependencies, and flows, showing that both payments and communication can be processed exclusively via the Service Hub.

With the Service Hub, two elementary products are to be distributed and made available in order to address the individual specifics of companies in terms of data integration. The core product is a basic version of software including several fundamental functions of a middleware for data integration. In addition, individual yet standardized modules for data integration are offered to adapt the basic version to the company-specific requirements. Based on the results of the literature research and the expert interviews, a subscription for the basic version seems suitable. Depending on whether the basic version of the software is easy to procure and does not need to be specially implemented at the customer' site, purchasing and procurement can be operated directly via the Service Hub. In the case of more complex implementation, an on-site implementation is also possible as an additional service from the software vendor. Offering these modules with a pay-per-use payment model might be appropriate, according to the expert interviews. The basic version and the modules together then result in the customer's individual version.

The platform provider, being one of the central elements of digital platforms, is implicitly represented by the platform itself and the software vendor, who at the same time also be a complementor. Thus, the integration of additional software vendors as complementors is an option for integrating third parties in the future platform design. Lastly, there is the customer represented by an industrial company that intends to purchase the corresponding solutions for data integration. By dividing the platform into a basic version and additional modules, the platform provider can benefit from future growth opportunities of the platform and at the same time profit from the growth and the need for data integration solutions of its customers [17]. The individual software version should be available both cloud-based and on premise, so that the customer can choose freely. The cloud infrastructure can be operated by the software vendor or sourced externally.

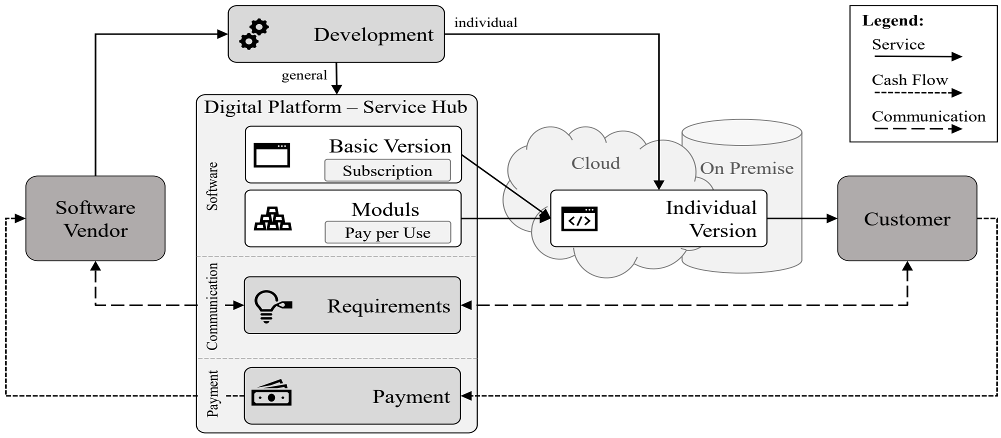


Figure 1: Design of the digital platform concept from a business model perspective (Service Hub)

Referring once again to the basic idea of an app store, one central difference can be noted. Unlike an app store, our various modules interact with each other and are therefore coordinated. In this solution, these modules can be used both on premise and in a cloud, which is often different with an app store. Thus, our approach complements the in literature identified gap of missing data integration solutions in the middleware area.

For the operation of the Service Hub, the following recommendations for action can be derived. In order to make the use of the Service Hub as convenient and simple as possible while still maintaining customer contact, it makes sense to design the payment flow as well as the communication between the customer and the software vendor directly via the platform. For communication, a forum, contact forms and other common methods may be implemented, analogous to conventional platforms. Via these interfaces needs, requirements and requests can be effectively communicated and documented. Analysis and evaluation of the correspondence will enable the platform provider to better understand the customer and facilitate continuous development of the platform. It supports continuous expansion of the range of available modules and adaption to the specific needs of the customer. However, since not all modules required in the long term might be available when Service Hub launches, the platform provider should focus initially on the most prioritized modules and then strategically expand the portfolio of modules as required.

4. Discussion and Limitation

The developed concept of the Service Hub offers several advantages and disadvantages as well as various future design options. A crucial advantage for the provider results from the marketing of standardized modules, which avoid new developments and the implementation for individual requests of data integration solutions, leading to a significant reduction of effort and costs [29]. Furthermore, the platform concept offers the possibility of flexible growth both with existing customers and new customers by improving the modules and adding new modules. Depending on the design of the platform, mixed forms of distribution are possible, in which the basic version or specific complex modules are installed in the traditional way and the platform is only used to offer further modules. In some cases, this can be technically easier to implement. In addition, it can be advantageous to expand the range of data integration options by opening the Service Hub to software solutions and modules from third-party providers to may increase the attractiveness of the Service Hub and to make it possible to profit financially from third-party providers by brokering their offerings. One advantage for the customer is the new convenience of purchasing data integration solutions. For an implementation or extension of the services, it is no longer necessary to contact the software provider for every single addition. Customers can conveniently make their adjustments according to the self-service system. In contrast, however, there are also drawbacks for both sides. Primarily, the direct, personal customer contact is reduced, which leads to disadvantages such as the software provider possibly losing sight of the customer's needs, which can affect the relevance of the product. It is therefore important to keep customer contact and communication as high as possible, both via the platform and in person. Additionally, there is a financial risk for the software provider, as the initial construction and implementation of the platform can lead to higher costs than a classic implementation, which may only pay off with a high number of customers in the future.

Besides advantages and disadvantages of our platform concept itself, our study naturally disposes of limitations and prospects for further research. First, there are limitations with regard to an industry-wide representativeness of our study, as we only conducted a sample of 10 expert interviews in the industry. Second, we did not consider the technical feasibility of our developed digital platform concept in detail and thus did not consider possible technical limitations as well as the trade-off between possible implementation costs and the benefits of our concept. Further research could address these limitations and include interviewing more experts and analysing the trade-off between the costs of technical implementation and the benefits of the concept in a business case. When compared to existing approaches, the Service Hub is to be understood as an integrative approach that allows existing technical solutions to be offered to the customer. Summarizing, the contribution of our approach is rather to be classified as an exaptation than a ground-breaking invention, since we apply known solutions to a new problem.

5. Conclusion

Due to various drivers and the resulting trends, it is likely that data integration will become increasingly important in manufacturing companies in the future. In our study, we therefore developed a conceptual model of a digital platform - the Service Hub - based on literature research and expert interviews. The Service Hub supports companies in their digital transformation by offering (data) integration solutions via individualized marketing of frequently used data (integration) processes. Despite the limitations outlined in this study, we have demonstrated a viable concept to easily commercialize various data integration processes that can help software and middleware vendors improve their business model.

References

[1] 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, 1–14.

[2] Reinhart, G., Bank, L., Brugger, M., Hieronymus, A., Köberlein, J., Roth, S., Bauernhansl, T., Sauer, A., Bauer, D., Kaymakci, C., Schel, D., Schlereth, A., Fridgen, G., Buhl, H.U., Bojung, C., Schott, P., Weibelzahl, M., Wenninger, S., Weigold, M., Lindner, M., Ronge, K., Oeder, A., Schimmelpfennig, J., Winter, C., Jarke, M., Ahrens, R., 2020. Konzept der Energiesynchronisationsplattform: Diskussionspapier V3.

[3] Kaymakci, C., Wenninger, S., Sauer, A., 2021. A Holistic Framework for AI Systems in Industrial Applications. Wirtschaftsinformatik 2021 Proceedings.

[4] Bauer, D., Stock, D., Bauernhansl, T., 2017. Movement Towards Service-orientation and App-orientation in Manufacturing IT. Procedia CIRP 62, 199–204.

[5] 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. Universität Augsburg; MDPI AG, Augsburg, Online-Ressource.

[6] Schott, P., Ahrens, R., Bauer, D., Hering, F., Keller, R., Pullmann, J., Schel, D., Schimmelpfennig, J., Simon, P., Weber, T., Abele, E., Bauernhansl, T., Fridgen, G., Jarke, M., Reinhart, G., 2018. Flexible IT platform for synchronizing energy demands with volatile markets. it - Information Technology 60 (3), 155–164.

[7] Gregor, Shirley, Alan R. Hevner, 2013. Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37 (2), 337–355.

[8] Hevner, A., March, S.T., Park, J., Ram, S., 2004. Design Science in Information Systems Research. MIS Quaterly (28 (1)), 75–105.

[9] Berger, S., Bitzer, M., Häckel, B., Voit, C., 2020. Approaching Digital Transformation - Development of a Multi-dimensional Maturity Model. ECIS 2020 Research Papers.

[10] Gimpel, H., Röglinger, M., 2015. Digital Transformation: Changes and Chances – Insights based on an Empirical Study, Project Group Business and Information Systems Engineering (BISE) of the.

[11] Siepermann, M., 2018. Predictive Analytics. https://wirtschaftslexikon.gabler.de/definition/predictive-analytics-54501/version-277530. Accessed 5 January 2021.

[12] Sanderson, D., Chaplin, J.C., Ratchev, S., 2020. Affordable Data Integration Approach for Production Enterprises. Procedia CIRP 93, 616–621.

[13] Cereda, P.R.M., Neto, J.J., 2017. A middleware architecture for adaptive devices. Procedia Computer Science 109, 1158–1163.

[14] Mertens, P., Bodendorf, F., König, W., Schumann, M., Hess, T., Buxmann, P. (Eds.), 2017. Grundzüge der Wirtschaftsinformatik. Springer Berlin Heidelberg, Berlin, Heidelberg.

[15] Bourhis, P., Reutter, J.L., Vrgoč, D., 2020. JSON: Data model and query languages. Information Systems 89, 101478.

[16] Schuh, G., Prote, J.-P., Gützlaff, A., Thomas, K., Sauermann, F., Rodemann, N., 2019. Internet of Production: Rethinking production management, in: Wulfsberg, J.P., Hintze, W., Behrens, B.-A. (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 533–542.

[17] Hein, A., Schreieck, M., Riasanow, T., Setzke, D.S., Wiesche, M., Böhm, M., Krcmar, H., 2020. Digital platform ecosystems. Electron Markets 30 (1), 87–98.

[18] Evans, D.S., Schmalensee, R., 2012. The Antitrust Analysis of Multi-Sided Platform Businesses: Law & Economics Working Papers. University of Chicago Institute for Law & Economics Olin Research Paper (No. 623).

[19] Jovanovic, M., Sjödin, D., Parida, V., 2021. Co-evolution of platform architecture, platform services, and platform governance: Expanding the platform value of industrial digital platforms. Technovation, 102218.

[20] Foerderer, J., Kude, T., Schuetz, S.W., Heinzl, A., 2019. Knowledge boundaries in enterprise software platform development: Antecedents and consequences for platform governance. Info Systems J 29 (1), 119–144.

[21] Constantinides, P., Henfridsson, O., Parker, G.G., 2018. Introduction—Platforms and Infrastructures in the Digital Age. Information Systems Research 29 (2), 381–400.

[22] Hein, A., Schreieck, M., Wiesche, M., Böhm, M., Krcmar, H., 2019. The emergence of native multi-sided platforms and their influence on incumbents. Electron Markets 29 (4), 631–647.

[23] Song, P., Xue, L., Rai, A., Zhang, C., 2018. The Ecosystem of Software Platform: A Study of Asymmetric Cross-Side Network Effects and Platform Governance. MISQ 42 (1), 121–142.

[24] Mesbahi, M.R., Rahmani, A.M., Hosseinzadeh, M., 2018. Reliability and high availability in cloud computing environments: a reference roadmap. Hum. Cent. Comput. Inf. Sci. 8.

[25] Übelhör, J., 2019. Industrieunternehmen und die Transformation von Geschäftsmodellen im Kontext der Digitalisierung – Eine empirische Studie über die Auswirkungen anhand des Business Model Canvas. HMD 56 (2), 453–467.

[26] Sousa-Zomer, T.T., Magalhães, L., Zancul, E., Cauchick-Miguel, P.A., 2018. Exploring the challenges for circular business implementation in manufacturing companies: An empirical investigation of a pay-per-use service provider. Resources, Conservation and Recycling 135, 3–13.

[27] Lipinski, K., Lackner, H., Laué, O.P., Kafka, G., Niemann, A., Raasch, E., Schoon, B., Radonic, A., 2019. Pay-per-Use. https://www.itwissen.info/Pay-per-Use-pay-per-use.html. Accessed 10 December 2020.

[28] Behrens, B.-A., Brosius, A., Hintze, W., Ihlenfeldt, S., Wulfsberg, J.P., 2021. Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg.

[29] Böhmann, T., Krcmar, H., 2004. Servicedatenmanagement für modulare Dienstleistungen, in: , Betriebliche Tertiarisierung : der ganzheitliche Wandel vom Produktionsbetrieb zum dienstleistenden Problemlöser. Dt. Univ.-Verl., Wiesbaden, pp. 149–178.

Biography

Julia Donnelly (*1994) is working at the FIM Research Center of the University of Augsburg and at the Project Group Business & Information Systems Engineering of the Fraunhofer FIT and focuses on research in the fields of digital transformation and digital business models.

Andreas John (*1995) is working as an IT consultant in the field of Industry 4.0 at soffico GmbH.

Jonas Mirlach (*2000) is studying Industrial Engineering at the University of Augsburg and is working at the FIM Research Center of the University of Augsburg with a research focus in the fields of digitalisation and IoT.

Kilian Osberghaus (*1999) is studying Industrial Engineering and Economics at the University of Augsburg and is working at the FIM Research Center of the University of Augsburg with a research focus in the field of digital health.

Silvia Rother (*1995) is studying Industrial Engineering at the University of Augsburg and is working in the process and quality management of the drive development at Hilti Entwicklungsgesellschaft mbH, Kaufering.

Christian Schmidt (*1991) is working as a Senior IT consultant and system architect for decentral communication patterns in the field of Industry 4.0 at soffico GmbH.

Hannes Voucko-Glockner (*1999) is studying Economics at the University of Augsburg and is working at the FIM Research Center of the University of Augsburg with a research focus in the field of data-based services in Industry 4.0 and smart and sustainable cities.

Simon Wenninger (*1995) 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 and focuses on research in the fields of data analytics in an industrial and energy context.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Extended Simulation Model For An Aerodynamic Feeding System

Torge Kolditz[1], Jakob Hentschel[1], Fabian Katz[1], Annika Raatz[1]
[1]Leibniz University Hannover, Institute of Assembly Technology, An der Universität 2, 30823 Garbsen, Germany

Abstract

Due to an increased number of product variants and shorter product life cycles, flexible automation plays a vital role in the producing industry. In assembly systems, industrial robots are used as highly versatile handling and joining devices. Simultaneously, the corresponding feeding systems that provide the workpieces in an orderly fashion for automated assembly can often not meet the required flexibility. In order to achieve high flexibility and reusability, an aerodynamic feeding system was developed. The feeding system can flexibly and rapidly adapt itself to new workpieces autonomously, using a genetic algorithm. To find the optimal parameters for the genetic algorithm, a workpiece specific simulation model of the aerodynamic orientation process was developed and validated in earlier work. In this work, we extended the simulation model with regard to the spectrum of workpieces that can be simulated and developed a user-friendly framework to simplify the application of the model. Our goal is to reduce the setting time of the genetic algorithm even further by predicting the optimal range of the feeding system's parameters for any workpiece using the extended simulation model. To evaluate and validate the simulation model, we carried out extensive tests with different exemplary workpieces. The results show that the setting time of the aerodynamic feeding system can be dramatically reduced using the extended simulation model, further increasing the flexibility and reusability of the system.

Keywords

Aerodynamic Feeding, Simulation, Flexible Feeding Systems

1. Introduction

Production in the modern industry is characterized by increasing numbers of product variants, shortening product life cycles and a high cost pressure due to global competition [1]. Especially in automated assembly systems, this results in reduced amortization periods and the demand for highly flexible systems [2]. Feeding technology, as one of the most complex and expensive parts of automated assembly systems [3], is therefore required to become more versatile. Commonly used solutions like vibratory bowl feeders cannot meet those demands [4]. Therefore, many approaches have been taken to increase the flexibility of conventional feeding systems, using interchangeable [5] or adjustable chicanes [6] [7]. At Leibniz University Hannover, an alternative solution for small part feeding was developed. The aerodynamic feeding system uses a constant jet of pressurized air to reorient workpieces with high speed and flexibility [1]. Due to the use of an air jet as chicane, the system is very flexible regarding the geometry of the workpieces. When changing between different workpieces, no hardware changes have to be made. Instead, the feeding system adjusts itself autonomously using a genetic algorithm [8]. Still, the initial values of the genetic algorithm are generated randomly within the system's parameter boundaries, which makes the expected setting time hard to predict. In this work, we present an approach to reduce the setting time and make it more predictable at the same time, using an extended simulation model of the aerodynamic orientation process. Before presenting our approach, we will introduce the aerodynamic feeding system used in our work.

2. Aerodynamic feeding system

The aerodynamic feeding system used in this work, manipulates the workpieces using an air jet with constant pressure. The principle is illustrated in Figure 1. The workpieces are first separated using a conventional centrifugal feeder, before they are handed over to a conveyor belt, which accelerates them to a desired velocity. Subsequently, the workpieces pass the nozzle, which produces an air jet, exercising a force on them. The extent of the force is thereby dependent on the geometry of the workpiece, in particular on the projected inflow area. If the projected inflow area variates along the longitudinal axis of the workpiece, a rotational impulse around the center of gravity is generated. For certain combinations of the five system parameters α, β, p, v and z (cf. Figure 1), workpieces arriving at the nozzle in one orientation are reoriented, while workpieces arriving in the other orientation are not reoriented.

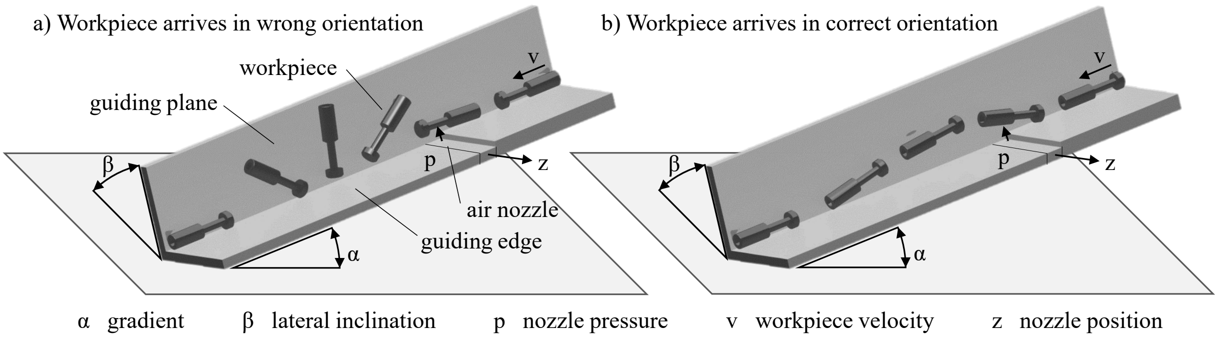


Figure 1: Principle of aerodynamic orientation

These parameter combinations are different for every type of workpiece and cannot easily be predicted. Therefore, a genetic algorithm that searches for a suitable parameter combination autonomously was implemented in the system control. In each generation, two new individuals are created. The parameter combinations of each individual are then set on the feeding system. Using a line scan camera, the orientation of the workpieces after the orientation process is determined. The ratio of correctly oriented workpieces to all the workpieces of one individual is defined as orientation rate and constitutes the fitness of each individual. For one individual, the feeding system tests 100 workpieces. In an iterative process, the individuals are evaluated, selected, recombined and mutated, until a satisfactory solution with an orientation rate of 95 % or higher is found. For a more detailed description please refer to [8] [9]. Using the genetic algorithm, the feeding system can adjust itself to new, unknown workpieces autonomously.

Still, in the current state, there are some disadvantages to the genetic algorithm. Firstly, the initial population of the genetic algorithm is created with random values in the parameter range of the aerodynamic feeding system. In combination with the heuristic character of the genetic algorithm, this can lead to varying and hardly predictable setting times. Secondly, the standard parameter range for the genetic algorithm, as shown in Table 1 is not optimal. In order to reach a high flexibility, the standard parameter range is set according to the system's mechanical, electrical or physical boundaries, regardless of the current workpiece properties.

Table 1: Parameter range of the aerodynamic feeding system for the genetic algorithm

Parameter	α	β	p	v	z
Minimum	20°	39°	0.01 bar	50 m/min	0 mm
Maximum	25°	50°	0.89 bar	80 m/min	10 mm
Increment	0.1°	0.1°	0.01 bar	1 m/min	1 mm

The genetic algorithm then searches the solution space spanned by these parameters with the exception of the nozzle pressure p. Using the extreme values of p, most workpieces would either not be lifted at all due to low pressure or shot out of the system due to high pressure. Therefore, the range for p has to be determined experimentally and manually before starting the genetic algorithm. For the other parameters, manual determination is not as easy and the standard range is used. As a result of this, the genetic algorithm potentially scans a lot of areas of the solution space, where no suitable solution for a particular workpiece can be found. This unnecessarily prolongs the setting time of the genetic algorithm. Furthermore, as aforementioned, the setting time of the genetic algorithm varies strongly and is hardly predictable, as can be seen in Table 2. In some test runs, a suitable solution was found after only testing two individuals of the genetic algorithm, while in other runs, up to 84 individuals had to be tested.

Table 2: Number of individuals tested before finding a suitable setting using the genetic algorithm in ten test runs

Test run No.	1	2	3	4	5	6	7	8	9	10
Tested individuals	36	11	66	2	4	84	23	13	2	6
Average (SD)	24.7 (±28.84)									

The aim of this work is to shorten the setting time of the aerodynamic feeding system for new workpieces, reduce the variations of the setting time and eliminate the need for experimental determination of the parameter range by the use of an extended, workpiece independent simulation model of the aerodynamic orientation process.

3. Extended simulation model

In this section, we will first briefly introduce the simulation model created in earlier work [8]. Then, we will present the extensions added to the model, like the implementation of a method to automatically calculate the workpiece properties (e.g. mass, inertia, center of gravity) as well as the development and validation of a method to calculate the lift of the workpiece created by the air jet (aerodynamic drag) under the consideration of the nozzle position z.

3.1 Existing simulation model

The original simulation model was created to assess the quality of different combinations of the system parameters α, β, p and v [8]. The parameter z was only added as system parameter later [9]. The Simulation was used to optimize the parameters of the genetic algorithm like population size, mutation rate or selection method. By doing so, the experimental effort was reduced drastically, while at the same time being able to perform thousands of virtual test runs in order to increase statistical confidence in the results.

In the simulation model, the workpiece is represented as a one-dimensional rod with the length, and inertia of the corresponding workpiece. The mass of the workpiece is represented as point mass positioned according to the center of gravity of the real workpiece. Since the workpiece's movement is restricted by the guiding edge and guiding plane (it is assumed that the workpiece is always in contact with the guiding plane), the equilibrium of forces (friction, aerodynamic drag, weight) and the equation of motion are only considered in two dimensions. The aerodynamic drag exerted on the workpiece by the air jet is calculated based on many parameters, like the distance of the workpiece to the nozzle, the nozzle pressure as well as the diameter and the drag coefficient of the workpiece. The resulting differential equations of second order, which will not be presented in detail here, cannot be solved analytically. Therefore, the simulation model is implemented and run in MATLAB/Simulink by Mathworks, where the equations are solved numerically. Comparisons between the simulated and the real trajectory of a workpiece show high conformity (see Figure 2) [8]. For better presentation, the one-dimensional rod is pictured as a CAD-Model of the workpiece.

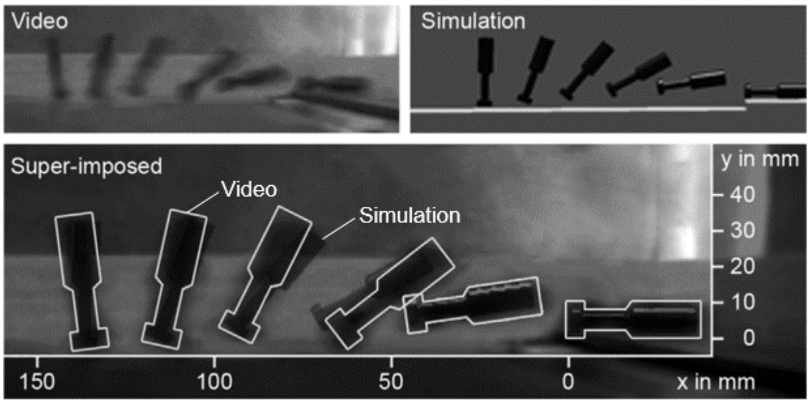


Figure 2: Comparison of real and simulated workpiece movement [8]

Nevertheless, a disadvantage of the simulation model is the low flexibility regarding the workpiece geometry. Busch et al. [8] designed the simulation model for a pneumatic plug, which can be seen as exemplary workpiece in Figure 1 and Figure 2. To simulate different workpieces, we would have to calculate all the workpiece properties for each different workpiece manually. Furthermore, the model only considers four system parameters. The nozzle position z, which we only implemented in the real system at a later time, is not considered in the existing model. Therefore, in the following sections, we will first describe the adaptations we implemented to enable the simulation model to calculate the movement of any rotationally symmetrical workpiece. Then, we will present our method to determine the aerodynamic drag of the air jet on the workpiece under consideration of different geometries and the nozzle position z.

3.1 Automatic determination of workpiece geometry and properties

As mentioned above, the simulation model calculates the movement of the workpiece based on the motion equations of a one-dimensional rod in the two-dimensional x-y-plane (cf. Figure 2). This reduces development and computational effort while still accurately representing the workpiece movement. When the workpiece geometry changes, certain properties of the rod, like the length, inertia and the center of gravity, change, but the underlying motion equations remain unchanged. In order to make the simulation model more universal and user-friendly, we enabled the model to determine the physical and geometric properties of a workpiece based on workpiece data provided by the user. The user provides the data by separating the workpiece into different sections that can be represented by primitive bodies, like cylinders or (truncated) cones. The model differentiates between outer and inner geometry features like bores. This way, various rotationally symmetrical parts can be represented very easily without the need for CAD-Software. In addition, separating the workpiece into sections that can be described with only few variables reduces the computational effort, shortening simulation time. With the dimensions known, the user only has to provide the density of each section and the model automatically determines all the necessary properties.

3.2 Calculation of aerodynamic drag under consideration of variable nozzle position

While the properties of a workpiece are constant during the entire simulation, the forces exerted on the workpiece change in dependence of the position, orientation, velocity and acceleration of the workpiece. Friction and inertial forces and moments can be calculated analytically, but the calculation of the aerodynamic drag of the workpiece in the pressurized air jet constitutes a particular challenge. In order to calculate the force F_w exerted on the workpiece by the air jet, we consider the air jet to be a cone, with the apex sitting in the nozzle. We calculate F_w using Equation 1:

$$F_w = c_w \cdot \frac{\rho}{2} \cdot w^2 \cdot A_{st} \qquad (1)$$

We assume the drag coefficient $c_w = 1.2$ for a cylinder and $\rho = 1.2041\ kg/m^3$ for air at ambient pressure. The air velocity w, with which the air jet hits the workpiece, is dependent on the distance between the workpiece surface and the nozzle as well as the nozzle pressure and diameter. The inflow area A_{st} is dependent on the workpiece geometry and the nozzle position z. In order to calculate A_{st}, we define an ellipse, representing a cross section through the air jet cone based on the distance between the workpiece surface and the nozzle. Still, the area of the ellipse is not necessarily the inflow area A_{st}, since it can exceed the dimensions of the workpiece in dependence of the relative position of the workpiece and the nozzle. In order to calculate the intersection of the ellipse and the workpiece accurately and efficiently, we chose a finite element approach (see Figure 3). An algorithm breaks down the workpiece and the ellipse into two, two-dimensional, binary matrices, where a “1” represents the cross-section of the workpiece or the ellipse respectively. All other entries are “0”. The workpiece matrix does not change during the simulation, while the matrix representing the ellipse changes in each time step, based on the relative position of the workpiece and the nozzle. When both matrices are added, the number of “2”s represents the inflow area A_{st} (white areas in Figure 3). Compared to other approaches (numerical integration, MATLAB polyshape function) our method showed much shorter computing times, while still achieving sufficient accuracy.

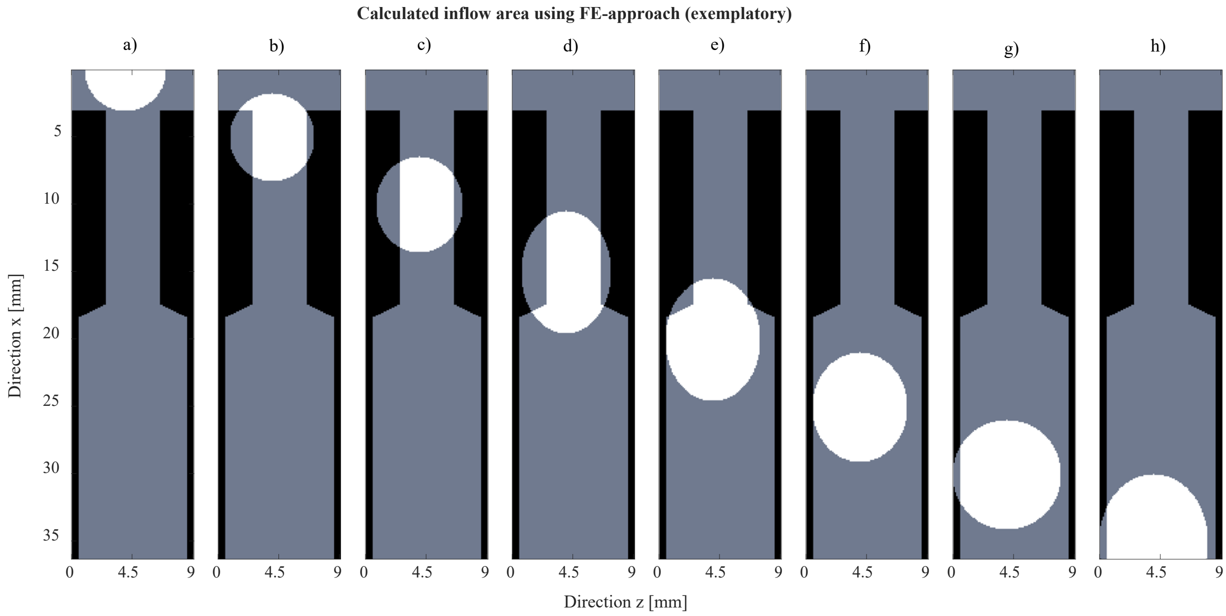


Figure 3: Calculated inflow area using FE-approach with a nozzle position of 4 mm in z-direction

In order to validate the approach, we compared the simulated aerodynamic drag force F_w with experimental measurements. For the validation, we mounted an exemplary workpiece (pneumatic plug) to a force sensor and measured the force exerted on the workpiece by an air jet coming from the nozzle with a relative pressure of 0.48 bar. Measurements were taken in increments of one millimeter along the longitudinal axis of the workpiece (x-direction in Figure 3), for different distances between workpiece and nozzle as well as for different nozzle positions z. The comparison between measured and simulated values, as shown in Figure 4, indicates good conformity. In order to match the fluctuations of the measured drag force, we added a white noise of 10 % of the calculated value in the simulation. Especially Figure 4 a) shows that the simulation can reproduce the influence of the workpiece geometry and the nozzle position on the aerodynamic drag.

Using the simulation model, we can now predict the movement of different, rotationally symmetrical workpieces in dependence of the five system parameters α, β, p, v and z. The aim of this work is to use the simulation model to reduce the setting time of the real feeding system by predicting and narrowing the boundaries of the system parameters. Nevertheless, doing this manually would mean a disproportionate effort. Therefore, we developed a framework, which automatically determines the boundaries of the system parameters based on the workpiece data provided by the user.

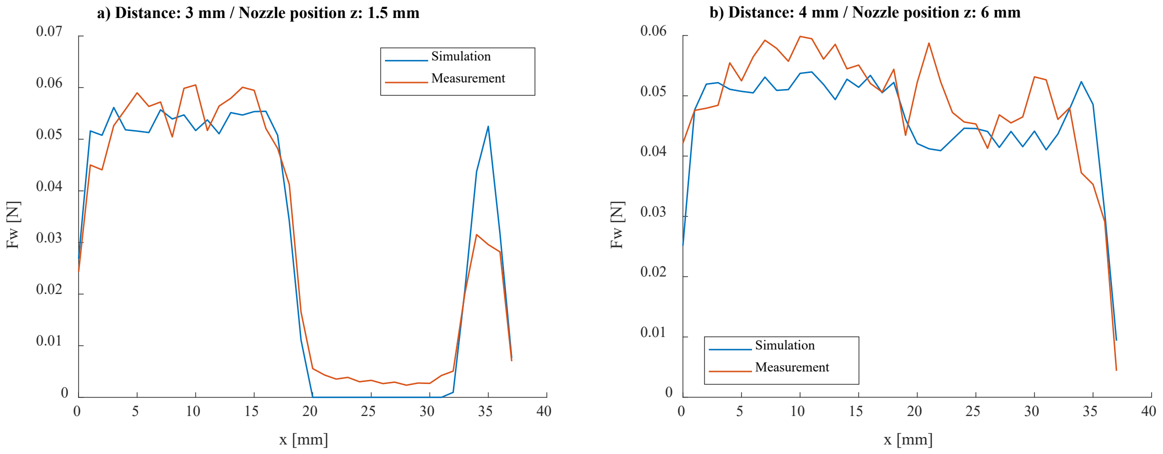


Figure 4: Comparison of aerodynamic drag (F_W) on an exemplary workpiece with nozzle pressure p = 0.48 bar for different workpiece positions relative to the nozzle

4. Framework for parameter prediction

In order to reduce the setting time of the aerodynamic feeding system, we integrated the optimized extended simulation into a new framework for an offline parameter prediction. The framework is implemented as a MATLAB script and directly interacts with the simulation. The main routine for parameter prediction contains two different subroutines, which systematically narrow down the lower, and upper bounds for each parameter. This approach drastically minimizes the search area for suitable parameter configurations of the genetic algorithm.

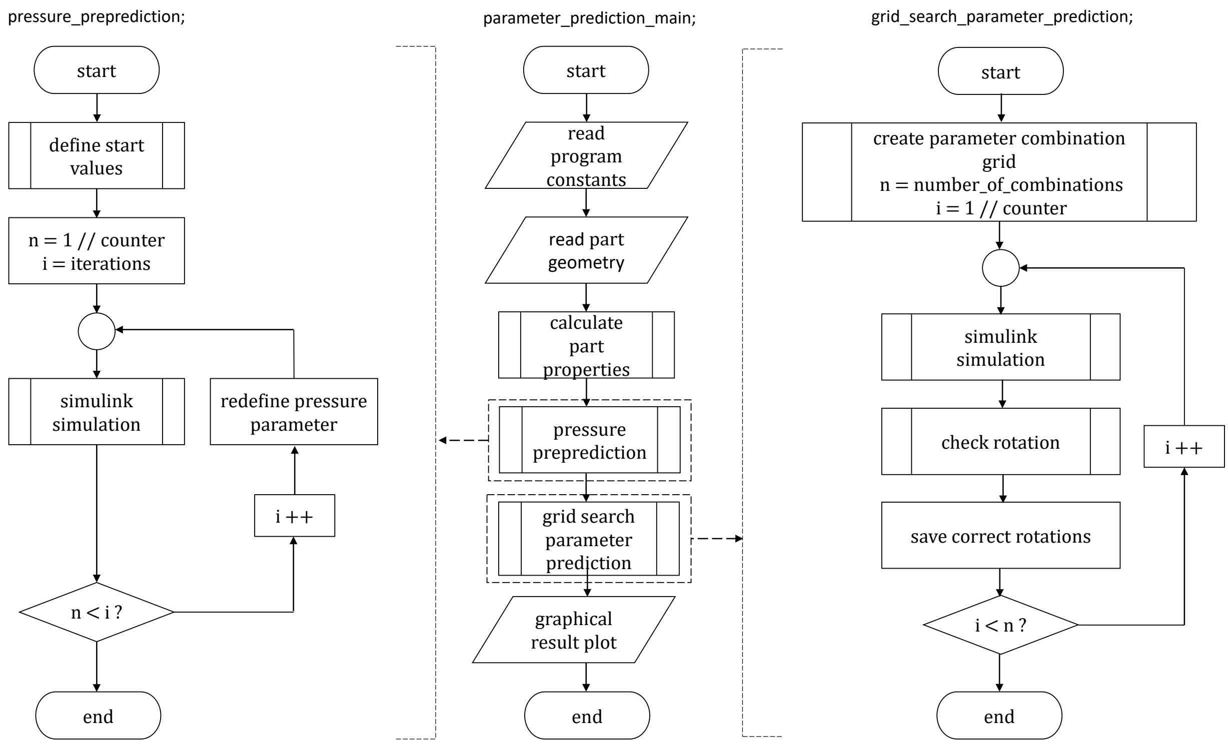


Figure 5: Flow chart of parameter prediction

Figure 5 shows the flow chart of the parameter prediction algorithm. After the initialization of the program constants, the main routine automatically reads and calculates the workpiece properties. Then, the main routine invokes two different subroutines. First, a pressure pre-prediction roughly narrows the pressure boundaries for the individual workpiece at hand. As mentioned above (cf. Section 2), the pressure range of the feeding system is much higher than the practical pressure range for an individual workpiece. In order to pre-predict the pressure, the upper and lower pressure boundaries are approximated numerically using the bisection method. The lower pressure boundary is defined as the lowest nozzle pressure to cause a reorientation in either direction; the upper pressure boundary is the lowest pressure to cause a double rotation of the workpiece in either direction. To predict the lower boundary, the algorithm first simulates the orientation process for the mean of the maximum and minimum pressure (cf. Table 1). If a reorientation occurs, the pressure is set to the mean of the previous value and the minimum pressure and the simulation is run again. This process is reiterated, until no reorientation occurs. Then, the step size is halved again but the search direction is reversed. Therefore, the algorithm oscillates around the lower boundary, until the stop criterion is fulfilled. The method works analogous for the upper boundary.

For the next step, we use a grid search algorithm to scan the solution space spanned by the boundaries of the system parameters. With all five system parameters divided evenly into n discrete values, we would have to calculate n^5 combinations. In order to reduce the computational effort, in this work we will only use three parameters as dynamic values and set the other two as constants. A sensitivity analysis showed that out of the five system parameters, the nozzle pressure p, the nozzle position z and the workpiece velocity v have the greatest impact on the orientation process. Therefore, we defined them as variable parameters inside the framework. The gradient α and lateral inclination β have a much smaller influence on the orientation process, therefore we defined them as constants with α = 25° and β = 45°. Considering only three parameters instead of five decreases the computing time drastically and results into a three dimensional grid with n^3 possible parameter configurations. All of these configurations are then entered in the extended simulation model to calculate the specific rotational behavior for each configuration for both input orientations of the workpiece. A configuration is marked as suitable when the workpiece only rotates in one input orientation and passes the nozzle in the other orientation without a rotation.

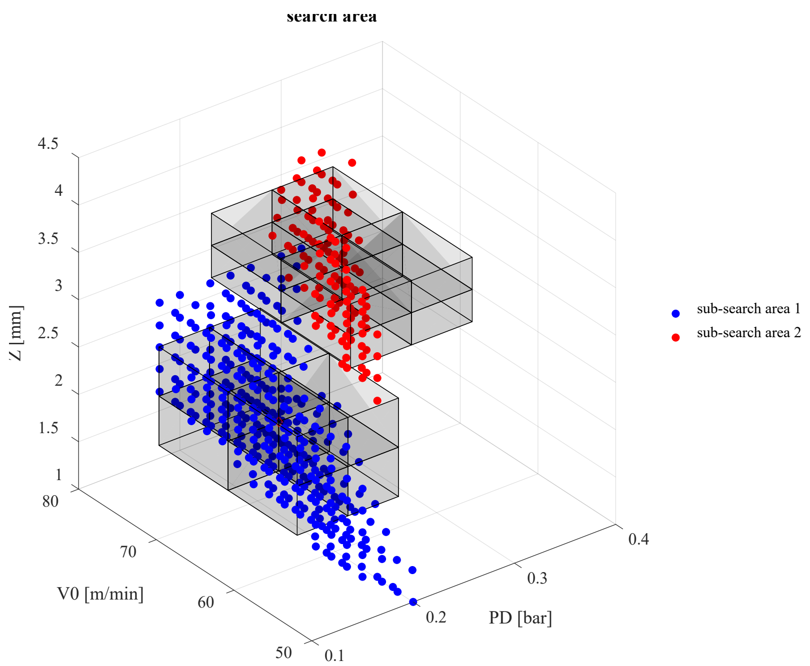


Figure 6: Suitable parameter configurations for exemplary workpiece (pneumatic plug) with two sub-search areas

Using a multi-layered graphical plot, the results for each parameter configuration are illustrated inside the three-dimensional configuration grid (see Figure 6). To determine which regions inside the search area are best suited, a density based scan function is used. The scan algorithm automatically distinguishes multiple sub-search areas with the highest density of suitable configurations and deletes statistical outliers. The distribution between the sub-search areas is visualized with different colors. For each sub-search area, the standard deviation for all three system parameters is plotted as boxes around the mean value configuration.

Figure 6 shows an example of the result plot for an exemplary workpiece, where two different sub-search areas were detected. A result summary gives a structured overview of the amount of correct rotations in each sub-search area as well as the mean values and standard deviation for each sub-search area. In this work, we used the mean and standard deviation of the sub-search area with the most correct orientations as predicted parameter boundaries for the genetic algorithm. In order not to narrow down the search space too much, we always rounded down the lower boundaries and always rounded up the upper boundaries.

5. Experimental evaluation

To experimentally evaluate the effect of the parameter prediction of the extended simulation model on the setting time of the aerodynamic feeding system, we compared the setting times of the genetic algorithm for three rotationally symmetrical workpieces (see Figure 7). One of the workpieces is an injection molded pneumatic plug (W1); workpieces W2 and W3 are self-constructed and printed using a resin 3D printer. For the evaluation, we first determined the average setting time of the system without the use of the simulation based framework and then repeated the experiments with the predicted parameter ranges.

Figure 7: Workpieces used for the evaluation of the parameter prediction framework

For the manual evaluation, the pressure boundaries are determined manually for the different workpieces and the corresponding duration for this process is measured. For the workpiece velocity v and the nozzle position z, the system's boundaries are selected. The gradient α is set to 25° and the lateral inclination β is set to 45°. Then, ten test runs are carried out for each workpiece. For each test run, the number of tested individuals is determined, as it is proportional to the setting time. After all test runs, the average number of tested individuals as well as the standard deviation is calculated.

For the evaluation of the simulation framework, the predicted boundaries for p, v, and z are implemented as boundaries for the genetic algorithm. The parameters α and β are also set to 25° and 45°. The start population of the genetic algorithm is no longer generated randomly but instead, for each dynamic parameter, the mean values of the predicted solution space are used. If a mean value lies between two adjustable values on the system, it is rounded up once and rounded down once, since the start population of the genetic algorithm has the size of two individuals. Subsequently, ten trials are carried out in order to compare average number of individuals tested by the genetic algorithm (setting time) and whether the duration per workpiece varies more or less in comparison to the manual values.

6. Results

The experiments show that the average setting time as well as the variation of the setting time can be drastically reduced with the use of the parameter prediction framework. Figure 8 shows the results of the ten test runs for each workpiece. Looking at the results of workpiece W1, we can see that the setting time was generally very high with the manually determined parameter boundaries. Also, the setting time varies very strongly between only two individuals at minimum and 84 tested individuals at the maximum. Considering that for each individual 100 workpieces run through the system, 8,400 workpieces have to be tested, before the feeding system reaches a satisfactory orientation rate. In addition, the determination of the boundaries of the nozzle pressure p took an experienced operator ten minutes on the feeding system. Using the simulation based parameter prediction framework, this additional preparation time is omitted. Furthermore, the parameter boundaries predicted by the framework lead to a much shorter setting time with less variation. Looking at the results of workpiece W2, we observed that even though the setting time is much shorter than for W1, the parameter prediction still cut the average number of tested individuals as well as the standard deviation in half. Here, the determination of the boundaries of p took eight minutes. The results of W3 are remarkable. While the genetic algorithm took an average of 20.5 individuals to find a satisfactory solution with the manual boundaries and a random start population, the start population provided by the framework produced satisfactory high orientation rates at the first tested individual in each of the ten test runs. The determination of the boundaries of the nozzle pressure took six minutes. This shows the potential of the simulation based parameter prediction framework to reduce the retooling time of the aerodynamic feeding system, by eliminating the need for manual preparatory work and drastically reducing the setting time.

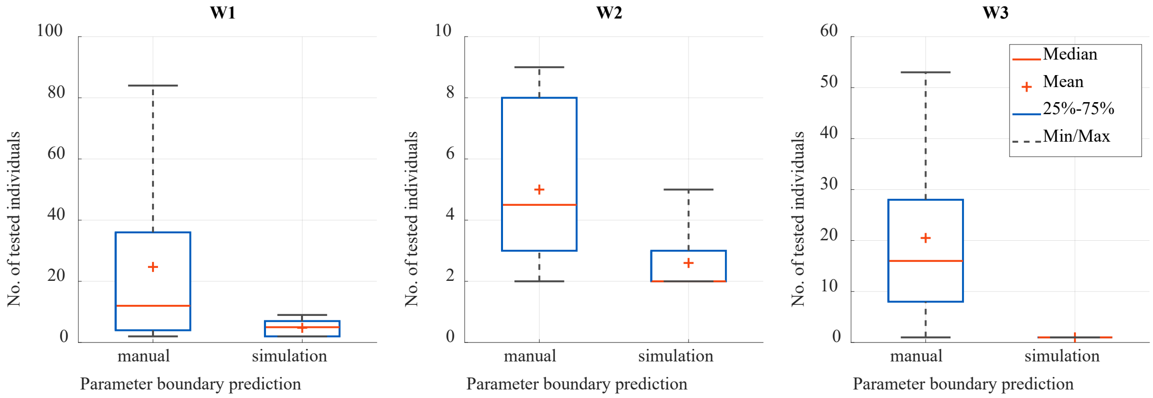


Figure 8: Number of tested individuals needed to reach an orientation rate of 95 % for different workpieces with manual and simulation based parameter boundary prediction

7. Conclusion and outlook

The experiments carried out with three simple exemplary components show very promising results regarding the reduction of the setting time, variation of the setting time and manual labor. For the three different workpieces, we reduced the number of parameter settings that had to be tested by the feeding system by 80, 50 and 95 %. We also reduced the standard deviation of the number of tested parameter settings of ten test runs by 90 and 60 % for the workpieces W1 and W2 respectively. Furthermore, a manual pre-prediction of the nozzle pressure range is no longer necessary. We therefore conclude that the parameter boundary prediction based on our extended simulation model is capable of drastically reducing the setting time and effort and increasing the flexibility of the aerodynamic feeding system.

In future work we will further extend the simulation model in order to increase the complexity of the workpieces that can be simulated and the precision of the parameter prediction. Preliminary experiments with more complex workpieces show that the parameter prediction framework does not always predict the best parameter boundaries. When the solution space determined by the simulation model is not as densely clustered as with the workpieces tested in this work, the mean and standard deviation do not accurately represent practical parameter boundaries. In order to resolve this issue, one aspect of future work will be the implementation of dynamic parameter boundaries for the genetic algorithm. In the current state, interrelations of the parameters are not considered; therefore, the solution space is always cuboid (assuming a solution space spanned by three parameters. Taking into account interactions between the parameters, the solution space determined by the simulation could be represented more accurately.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG - German Research Foundation) within the project "Model-based increase of the flexibility and robustness of an aerodynamic feeding system for high-performance assembly" (project number: 243351293).

References

[1] K. Knüppel und P. Nyhuis, „Aerodynamic part feeding technology - Flexible high-speed part provision for mass production," *Advanced Materials Research,* Nr. 907, pp. 165-179, 9 April 2014.

[2] L. Barbazza, M. Faccio, F. Oscari und G. Rosati, „Agilit in assembly systems: a comparison model," *Assembly Automation,* pp. 411-421, 7 März 2017.

[3] J. Krüger, T. Lien und A. Verl, „Cooperation of human and machines in assembly lines," *CIRP Annals,* Bd. Volume 58, Nr. Issue 2, pp. 628-646, 2009.

[4] S. Mathiesen, L. Sörensen, D. Kraft und L.-P. Ellekilde, „Optimisation of Trap Design for Vibratory Bowl Feeders," *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA),* pp. 3467-3474, 21-25 Mai 2018.

[5] M. Loy und G. Reinhart, „Design of a modular feeder for optimal operating performance," *CIRP Journal of Manufacturing Science and Technology,* Bd. Volume 3, Nr. Issue 3, pp. 191-195, 2010.

[6] A. Joneja und N. Lee, „Automated Configuration of Parametric Feeding Tools for Mass Customization," *Computers & Industrial Engineering,* Bd. 35, Nr. 3-4, pp. 463-466, 1998.

[7] T. Zhang, G. Smith, R.-P. Berretty, M. Overmars und K. Goldberg, „The Toppling Graph: Designing Pin Sequences for Part Feeding," *Proceedings of the 2000 IEEE International Conference on Robotics & Automation,* pp. 139-146, April 2000.

[8] J. Busch, M. Quirico, L. Richter, M. Schmidt, A. Raatz und P. Nyhuis, „A genetic algorithm for a self-learning parameterization of an aerodynamic part feeding system for high-speed assembly," *CIRP Annals - Manufacturing Technologiy,* Bd. 64, Nr. 1, pp. 5-8, 2015.

[9] T. Kolditz, C.-V. Ince und A. Raatz, „Investigation on the Convergence of the Genetic Algorithm of an Aerodynamic Feeding System due to the Enlargement of the Solution Space," *Proceedings of the 9th International Precision Assembly Seminar (IPAS),* 2021.

Biography

Torge Kolditz (*1993) has been a research associate at the Institute of Assembly Technology (*match*) since his graduation in mechanical engineering in 2018, where he researches in the field of automated assembly.

Jakob Hentschel (*1997) studies mechanical engineering and looks forward to graduate with a bachelor's degree in 2021. Since 2019, he has been working as a student research assistant at *match*.

Fabian Katz (*1994) studies mechanical engineering for his master's degree. He has been working as a student research assistant at *match* since completing his bachelor's degree in 2020.

Annika Raatz (*1971) has been the head of *match* since its establishment in 2013. Prof. Dr.-Ing. Annika Raatz is a member of the Society for Assembly, Handling and Industrial Robotics (MHI) and the Academic Association for Production Technology (WGP). Her main research topics are robot aided handling processes, machine concepts for handling, assembly and production automation and soft material robotic systems.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

A Deep Learning Framework for Automated Collection and Analysis of Traffic Data based on Identifying and Classifying Delivery Vehicles in Logistics

Meng Jin[1], Lars Andreas Mauch[2], Bernd Bienzeisler[1]

[1]*Fraunhofer IAO | Forschungs- und Innovationszentrum Kognitive Dienstleistungssystem , Heilbronn, Germany*
[2]*Fraunhofer IAO | Energy Innovation, Stuttgart, Germany*

Abstract

In urban logistics, analyzing urban traffic data plays an important role in achieving higher schedule reliability and delivery time efficiency. To increase the diversity of urban traffic data, we developed a solution for the automated collection and analysis of different types of traffic data. These are needed to optimize and control the flow of traffic.

The use of traditional on-road sensors (e.g., inductive loops) for collecting data is necessary but not currently sufficient because it cannot draw any conclusions about the type of goods being transported. In this paper, we propose a framework in which different classes of delivery vehicles and types of goods being shipped are identified in road videos by deep-learning-based image recognition method. Video sequences are automatically evaluated according to the following criteria: (i) distinguish between individual and commercial vehicles, (ii) identify the category of commercial vehicles, for example, van, box trucks, small trucks, etc. (iii) identify the special features of the vehicle body (such as the name of the carrier) to classify commercial transportation of food, general goods or package services, etc. Using this method, logistics throughput of a designated city or region and the peak time of goods transportation can be obtained. This provides the carrier with better pre-advice and potential actions to improve transportation efficiency. For the evaluation of our framework, we collected real street videos at different time points in the main traffic arteries of Heilbronn, Germany. In particular, the difference between traffic flow of logistics services before and during the COVID-19 epidemic was compared. The results of implementation and testing demonstrated a high-precision, low-latency performance of the framework for obtaining urban logistics data.

Keywords

Urban Freight Transportation; Identification and Classification Delivery Vehicles; Logistics; Deep Learning; Image Recognition

1. Introduction

The complexity of urban traffic and the diversity of its impacts require very thorough data collection methods since the collection and analysis of traffic data play a significant role in traffic safety, traffic pollution, and urban road planning. In this paper, we focus on the collection method of commercial vehicle traffic data, because they account for a large part of the traffic data and are particularly useful for urban logistics service providers, who compete with other road users for scarce traffic space in the inner city, the continuous traffic

DOI: https://doi.org/10.15488/11265 publish-Ing.

data are extremely important for them to optimize transportation operations and make reliable decisions through data analysis.

The technology of collecting traffic data has been studied for decades. It can be easily divided into two categories, the intrusive and non-intrusive methods [1]. The intrusive methods consist of a data recorder or a sensor placing on or in the road, such as traditional road sensors (induction loops). Such technologies have been employed for many years. However, due to various factors, including limited coverage and possible inability to obtain continuous traffic count data, such as pavement repair, construction, and maintenance. In this case, non-intrusive data can be used as an alternative for a short period or to supplement traffic data diversity, which is based on remote observation, including manual counting, passive and active infrared, microwave radar, video image detection, etc. [1]. These different types of data collectors have their advantages and disadvantages, and their combination provides more complete traffic data. In this paper, we mainly focus on video image detection technology because other data collection methods have made outstanding contributions to vehicle counting and geographic information. Obtaining more vehicle feature information through images, such as vehicle classification, vehicle size, weight, or the type of goods being transported, to supplement the defects of current traffic data is the focus of our work. According to literature surveys, most existing video-based detectors can only provide a few macro traffic parameters, such as traffic count and average speed. Several algorithms based on computer vision have been developed to classify vehicles, but most of them only use traditional image processing and mathematical methods, such as calculating the vehicle's length in pixels for classification [5]. The result is that trucks can only be divided into long vehicles and short vehicles, and the accuracy of the results cannot achieve very high since the vehicle's length is susceptible to road reflections, lighting changes, shadows, and other interference factors. In addition, manual input of the exact road area to be observed and the threshold used to determine the length of the truck are also required. In this paper, a new video-based traffic data collection framework is developed to track each passing vehicle, the types of vehicles and the features of vehicle body will be recognized using deep learning methods. It no longer relies on manual input of thresholds and precise road positions in advance. By identifying features on the vehicle, such as letters, logistics service providers and the items being transported can also be deduced. In this way, high-precision and detailed traffic data can be collected, and the process of data collection is simplified compared with the complex algorithm that required manual input before.

The paper starts with a short overview of the methods for collecting traffic data in previous works. In section 3, the emphasis is put on the framework for identifying the features of commercial vehicles on the road based on the deep-learning method, the detailed methods used to distinguish between individuals and commercial vehicles, classify commercial vehicles, and identify transported items will also be described in detail. The purpose of section 4 is to evaluate our proposed method and show the results of applying it to roads to collect traffic data. The final section concludes this research effort and proposes further research topics.

2. Related Work

Before presenting the details of our framework related studies are briefly introduced. The methods of traffic data collection have been an area of interest in intelligent transportation system for the past few decades. [1] introduced all data collection methods on the market and analyzed their advantages and disadvantages, such as capabilities and restrictions. [2] in 2010 provided a framework for optimizing logistics time based on traffic data collection and proved that traffic data is beneficial for estimating and optimizing logistics delivery time. [3] designed a traffic information system, as a test, traffic flow and vehicle speed were collected in Berlin, Germany through fixed measuring equipment and specially equipped vehicles. [4] analyzed the advantages and disadvantages of traffic data based on video detection technology and Automatic Traffic Recorder. The results showed that although the data based on video detection technology are not very accurate, they can provide useful traffic data information. The acceptability of the data should depend on the accuracy requirements of the use cases. [6] developed a new video-based traffic data collection system. Each passing vehicle can be tracked and classified in mixed traffic situations. The average speed of each passing vehicle is identified as the main contribution of this paper. [7] proposed a system based on video detection to classify vehicles and estimate the target color. [8] showed an algorithm for classifying vehicles based on five-dimensional feature vectors, and each vehicle can be tracked by a temporal tracking methodology. Through this process, both the macro and micro parameters of the vehicle will be available. However, due to the high computational complexity, real-time collection of traffic data is difficult to be applied. [9] proposed a method to detect and classify vehicles based on a three-dimensional (3D) model in aerial imagery, but the previous condition of this method is to build a 3D matching vehicle model for training, that is complicated and time-consuming, the accuracy is obviously affected by the constructed 3D model. Generally, these previous studies still have the following shortcomings: (i) many methods can only detect single-lane vehicles (ii) Manual input of the detection area or setting of detection parameters is required in advance (iii) Insufficient types of vehicles, such as only divided into two classes by length (iv) lacks an identification method to identify the transported object. The framework described in this paper aims to overcome these shortcomings.

3. Methodology

In this section, the main steps of our deep learning-based traffic data collection and analysis framework will be introduced, which can be briefly called DeepTraffic.

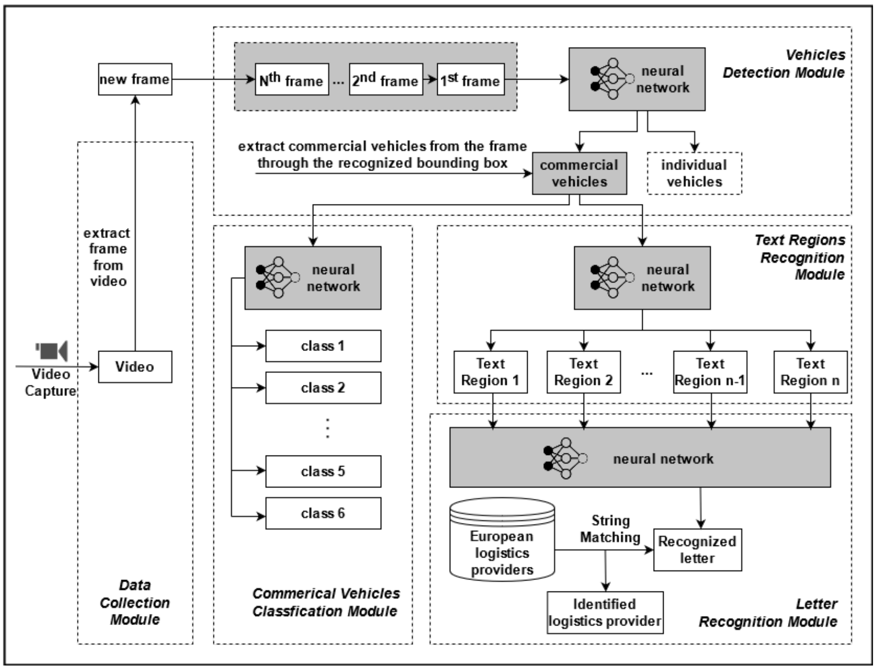


Figure 1: Flow Chart of the DeepTraffic Framework

3.1 Framework Overview

DeepTraffic framework has five modules: data collection, vehicle detection, commercial vehicle classification, text regions recognition, and letter recognition. Figure 1 shows the flow chart of our framework. First, a digital video is taken by a camera placed on the road to capture multiple lanes in the same direction. Our framework takes road video as input and extracts every frame in real-time as the new input for the next step. In the next vehicle detection module, we trained a neural network with various labeled real-world pictures to detect vehicles and distinguish between private and commercial vehicles. Through this step, the commercial vehicle can be marked with a bounding box in the video in real-time and extracted from the background as a new image, which is used as a new input and enters the following two modules at the same time: in the commercial vehicle classification module, we employed a simplified VGG network to divide commercial vehicles into six categories, in the text regions recognition module, an open-source deep learning model is utilized to obtain the features of the vehicle body. The recognized text area will be transferred to the letter recognition module, where all English or German letters can be identified from the picture. To ensure the correct rate of recognition, a string-matching algorithm is used to match the recognized letters with the name of the courier service provider in our pre-established database Finally, we tracked the vehicle and counted the number of passing vehicles by analyzing the vehicle's type and possible logistics providers in consecutive frames. The following sections will explain the methods involved in the entire process in more detail.

3.2 Collection of Image Samples

In a deep-learning-based framework, many sample pictures are needed to train the neural network. To make our system adapt to each road section, various video images on highways, country roads and urban roads in different periods and weather conditions are taken by *YI* cameras with a frequency of 30 frames per second (FPS). Figure 2 shows some examples of our images. We shot 6 hours of video on the bridge above highway A8 and B14 in Stuttgart, Germany, highway A6 in Heilbronn, Germany, and 9 hours of video along the main road in the city Heilbronn center. We collect data for research purposes only, according to data protection regulations we will not collect any personal information from the video. All video data will not be disclosed and will be deleted after analysis.

Figure 2: Examples of Training Images

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
1×	Convolutional	64	3 × 3	
1×	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
2×	Convolutional	128	3 × 3	
2×	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
8×	Convolutional	256	3 × 3	
8×	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
8×	Convolutional	512	3 × 3	
8×	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
4×	Convolutional	1024	3 × 3	
4×	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 3: YOLO V3 network structure [11]

3.3 Vehicle Detection

Deep Neural Networks (DNN) based vehicle detection methods have been applied to many image/video applications and have achieved state-of-the-art on various data set. A large data set is required for transfer learning if we apply them directly, which is a laborious and expensive construction. However, if we train them on a small data set to solve a complex task, it usually leads to overfitting problems. We intend to use DNN with our small data set to accomplish the following two simple goals: (i) detecting vehicles, (ii) distinguishing between private and commercial vehicles. The detailed vehicle classification network will be applied separately. Because if commercial vehicles are classified in detail on this complex DNN model, overfitting often occurs. After searching, we found that a DNN architecture called YOLO V3 [11] [12] has been widely used in object detection tasks and achieved excellent results. To perform vehicle detection more effectively, we choose the YOLO V3 framework as the basic structure and fine-tuned it based on our data set. More than 5,000 pictures from 25 hours of video are selected as a training image and labeled by an open-source software called LabelImg [10]. Figure 3 shows the basic structure of YOLO V3 network, which is trained on COCO dataset with 80 classes [11]. Therefore, we reduced the number of categories to 2 to suit our tasks, namely private and commercial vehicles. Figure 4 shows the detection results of the YOLO V3 model fine-tuned by our data set: some vehicles in the distance have been identified, but this is not what we expected because we need clear pictures to prepare for future vehicle classification or body feature recognition. We used the following criterion to remove these outliers: (i) we only select objects with a confidence value higher than 0.6 as valid objects. (ii) we set a general road section (the green box in Figure 4) as a valid region, and the detected object is only valid when the center of the object is within the selected valid region. The results show that the YOLO V3 model can identify all private and commercial vehicles in the video with an accuracy rate of 97%. Since we are analyzing consecutive frames, even if certain vehicles are not recognized in some individual frames, these vehicles will appear in at least three frames to ensure that no passing vehicles are missed.

Figure 4: Examples of detected Vehicles

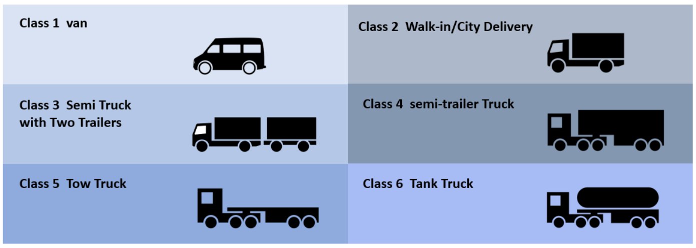


Figure 5: Examples of detected Vehicles

3.4 Commercial Vehicles Classification

In the previous step, the recognized commercial vehicle will be extracted from the video automatically and saved locally used as the new input for the vehicle classification module. The corresponding frame number is also stored in the image name and used later for vehicle counting. As shown in Figure 5, we classify vehicles into six classes: van, walk-in/city delivery, semi-truck with two trailers, semi-trailer truck, tow truck, and tank truck. These types of vehicles are commonly used in German logistics transportation. The CNN architecture we used for classification is SmallerVGGNet, a simplified version of the original VGGNet, which was first introduced in [13] in 2014. We do not need such a vast network structure of the original version for our task with only six categories, so we simplified the network with only four repeated Conv blocks with different kernel sizes (see Figure 6). The input image of the network is a fixed-size RGB image. We resized all the extracted images to 224*244 and sent them to the network. First, four conv blocks are used to extract features. Each conv block contains one 2D convolutional layer with different kernel sizes, a Relu activation function to scale the output, one batch normalization layer to standardize the inputs, and one max pooling layer for aggregating the results of convolutional layers by only passing on the strongest feature. The output should be rolled out through Flatten layer and connected to a dense layer, which has exactly the number of neurons that corresponds to the number of different classes to be recognized. In this case, we used 600 classified images as the data set, of which 480 are used as training images and 120 are used as test images. The network was trained on a Nvidia 2080Ti GPU, the accuracy of the training set reached 98.2%, and the accuracy of the test set fluctuated between 90% and 95% after 30 Epochs.

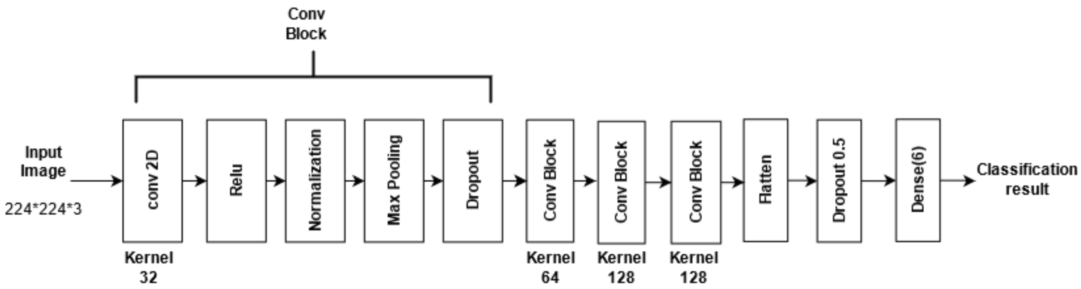


Figure 6: Structure of SmallerVGGNet

Figure 7: Example of text recognition on the vehicle body

3.5 Text Regions Recognition

To detect the text on the vehicle body, we directly used an open-source deep learning model called EAST proposed by Zhou et al. in 2017 [14]. It can run at near real-time at 13 FPS on 720p images and obtain state-of-the-art text detection accuracy. The vehicle image extracted in the vehicle detection module is also adopted as input in this module. During the recognition process, we found that the text on the car body is not parallel in the picture because of the shooting angle, as shown in the leftmost picture in Figure 7, which makes the text harder to recognize. To overcome this difficulty, the input picture is preprocessed: according to the angle at which we took the video, the picture can be rotated 5 degrees clockwise to make the text in the image parallel, as shown in the middle picture in Figure 7, this angle of rotation can be applied to almost all the pictures we took. The rightmost picture in Figure 7 shows the final recognition result. The text area on the vehicle body will be recognized and cut out.

3.6 Letter Recognition and Matching

To identify the logistics provider and assume the items being shipped, we need to convert the recognized text into letters. In this module, we used the open-source scene text recognition (STR) benchmark [15], which was proposed by Baek et al. in 2019 and integrated and analyzed the current advanced models for character recognition. It pointed that the accuracy of the best model on different data sets is between 74% and 94.9%, which means that using this model to recognize the characters in our cases cannot be completely accurate. Figure 8 shows a part of the recognition result, that is stored in a text file and contains the following information: the ID of the video, the number of frames in the video, the position of the recognized vehicle in the frame, the recognized vehicle type, the probability of the recognized vehicle type, and the characters recognized by the vehicle body. From the results, we concluded that some recognized characters are slightly different, even in the same vehicle in consecutive frames. To obtain the unique logistics provider, we collected information on 5188 logistics companies in Germany and Austria in advance and calculated the similarity between the recognized characters and them through the Levenshtein algorithm and matched it to the unique logistics provider in our database. In consideration of data protection, we do not public our logistics company database here. After matching the logistics provider in our database, we count the number of vehicles according to the following criteria: (i) if there are vehicles of the same logistics provider in consecutive frames, this vehicle can be regarded as one vehicle, and one vehicle is counted. (ii) if there are no characters recognized in consecutive frames, such as a white vehicle, then we use the position of the identified vehicle in the picture to determine whether it is the same vehicle, because as the car moves forward, the position of the car must be closer to our camera. To determine the type of vehicle, we perform the following calculation: In consecutive frames, if the identified vehicle type is unique, the vehicle type can be simply derived. If more than two different types are identified, the type with the highest probability is selected as the final result. The item being transported can be inferred from the name of the logical operator. For example, from the result of "DACHSER Intelligent logistics", it can be known that the item being transported is food, if "DHL or DPD" are recognized, it can be inferred to be a package service. An example of result of our framework is shown in Figure 9.

```
video11_frame0139_000_347_Semi Truck with two trailers_099_
video11_frame0188_061_320_Tow Truck_099_
video11_frame0196_296_462_Semi Truck with two trailers_099_Intelligent|DACHSER
video11_frame0197_260_450_Semi Truck with two trailers_099_Intelligent|DACHSER
video11_frame0198_213_431_Semi Truck with two trailers_099_DACHSER|Intelligent
video11_frame0199_158_424_Semi Truck with two trailers_099_Intelligent|DACHSER
video11_frame0200_105_392_Semi Truck with two trailers_099_Intelligent|Logistics|DACHSER
video11_frame0201_022_364_Semi Truck with two trailers_099_Intelligent|Antoristically|DACHSER
video11_frame0202_000_344_Semi Truck with two trailers_099_Logistics|STACHS|DACHSER|Intelligent
video11_frame0203_006_307_Semi Truck with two trailers_099_Logistics|Intelligent|DACHSER
```

Figure 8: Example of text recognition

```
['video10', 'frame0002', 'Walk-in', 'Concerverting']
['video10', 'frame0029', 'Tank Truck', 'ABE']
['video10', 'frame0062', 'Semi-trailer Truck', 'Wiedemeyer']
['video10', 'frame0069', 'Tow Truck', '']
['video10', 'frame0101', 'Semi-trailer Truck', 'EUROTIRE']
['video10', 'frame0131', 'Semi-trailer Truck', 'BOR']
['video10', 'frame0145', 'Semi-trailer Truck', 'KING']
['video11', 'frame0010', 'Semi-trailer Truck', 'burkhardt']
['video11', 'frame0022', 'Semi-trailer Truck', 'Simon SPEDITION Darmstadt']
['video11', 'frame0071', 'Semi-trailer Truck', 'ACTION'],
['video11', 'frame0081', 'Tank Truck','Alloborthors Contranter transter']
['video11', 'frame0115', 'Semi-trailer Truck', 'ACTION']
['video11', 'frame0203', 'Semi Truck with two trailers', 'DACHSER Inteligent Logistics']
```

Figure 9: Example of text recognition on the vehicle

4. Results

Table 1: Experimental results for vehicle detection

	total	True Positive	True Negative	False Negative	False Positive	correct (%)
result	630	612	10	0	9	97.1

Experimental results and discussion on the performance of the DeepTraffic framework are described in the section. To demonstrate the effectiveness of our framework, the experiments are conducted with the following procedures: we chose 9 hours of video data as the test set, which were shot on three main roads in Heilbronn, Germany. Considering the intensity of the sun and the weather conditions, videos are shot between 9:00-10:00, 13:00-14:00, and 16:00-17:00 on sunny, cloudy, and rainy days. No input images are included in the sample images for training. Table 1 shows the results of vehicle detection modules. We compared the results from our framework with the results observed manually (ground-truth data). Among the 630 commercial vehicles, 612 were correctly identified and the recognition accuracy rate reached 97.1%. 10 commercial vehicles were not identified because of occlusion, 8 large private vehicles, such as SUVs, were incorrectly identified as commercial vehicles.

Table 2: Experimental results for vehicle classification

	Class1	Class2	Class3	Class4	Class5	Class6	total	correct	%
Class1	65	18	0	1	0	0	84	65	77.3

Class2	0	218	0	16	0	1	234	218	93.2
Class3	0	0	28	4	0	0	32	28	87.5
Class4	0	3	2	231	0	0	236	231	97.9
Class5	0	0	0	2	24	1	27	24	88.9
Class6	0	0	0	3	0	14	17	14	82.3
total							630	580	92.1

*Class1=Van, Class2=Walk-in, Class3= Semi-Truck with two trailers, Class4= Semi-trailer truck, Class5=Two truck, Class6=Tank truck

Table 2 shows the results of our evaluation of one-hour video classification, which classifies the commercial vehicles in each frame. The correct average rate is close to 92.1%. In this experiment, only when the classification result matches the human visual judgment results, the system classification results are regarded as the correct answer. The later steps of using vehicle counting also bring more accurate classification results since the same vehicle may be divided into different categories due to different distances in consecutive frames. The final classification results will be analyzed through the subsequent steps of vehicle counting, and the category with a higher probability will be selected. It can also be concluded from the experimental that the results on cloudy and rainy days are better than those for sunny days because the images are more stable without the influence of sunlight on cloudy or rainy days.

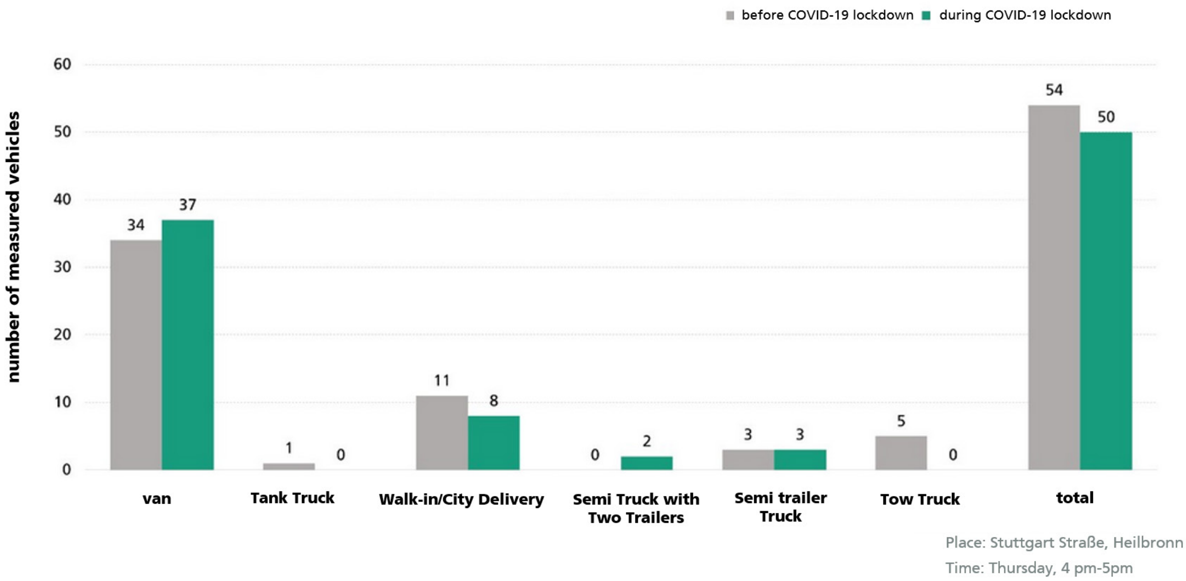


Figure 10: Traffic flows before and during the COVID-19 lockdown

In consideration of data protection, the result of the next text regions recognition module and letter recognition module will not be shown here. From the results, we can conclude that, although the text recognition algorithms have a high accuracy rate, only one-third of passing vehicles are clearly identified which logistics provider belongs to. There are three reasons why the logistic provider has not been identified: (i) there are no recognizable characters on the vehicle, (ii) there are only some characters used for advertising on the vehicle, which cannot match the providers in our database, (iii) the identified provider is not in our database. After analyzing the identified logistics providers, we found that the vehicles of German logistics providers appeared the most frequently. To determine the items they deliver, the service areas of them should be researched one by one. The second place is parcel services such as DHL, PDP, and GLS, which appear 1 to 5 times per hour. The third place is the delivery service of some supermarkets, which is used to deliver food and drinks.

To put our system into practice, we used our system to collect and process traffic data in the city of Heilbronn to understand whether lockdown has an impact on commercial traffic. We took two one-hour videos from 4 pm to 5 pm on Thursday at the same place in Heilbronn before (27.02.2020) and during lockdown (09.04.2020). Comparing the results, we counted that the amount of commercial traffic on Stuttgarter Street was almost the same as the traffic collected before the lockdown (see Figure 10). the results indicate that city delivery traffic was less affected by lockdown than private transport. It is noticeable that no tow truck was detected during the Corona period, obviously, there was less needed to tow away incorrectly parked or defective cars.

5. Conclusion and Outlook

This paper proposed a framework that combines different deep learning algorithms to collect and analyze each commercial vehicle on the road by using video images. The evaluation results from the three test locations in Heilbronn are encouraging. In all the test data, the accuracy of vehicle detection is higher than 97.1%, and the accuracy of vehicle classification and vehicle counting is higher than 92.1%. Letter recognition can identify the providers of one-third of the vehicles on the road. Of course, it depends to a large extent on the data in our pre-set database and whether the characteristics of the vehicle body are apparent. The test results show that the proposed DeepTraffic framework works stably and effectively under tested traffic conditions. However, the following problems still need us to improve further: (i) the current prototype framework will be affected by the environment: for example, it cannot work at night because our camera does not support taking night photos. Or during a rainy day, the text on the car body will be deformed due to rain, resulting in a decrease in the accuracy of recognition. (ii) vehicle occlusion significantly affects the accuracy of classification. In addition, more robust algorithms should be researched and explored to solve the light reflection problem to improve the reliability of the system.

References

[1] Leduc, G. (2008) 'Road traffic data: Collection methods and applications', Working Papers on Energy, Transport and Climate Change, 1(55), 1-55.

[2] Ehmke, J. F., Mattfled, D. (2010) 'Data allocation and application for time-dependent vehicle routing in city logistics', European Transport

[3] Fleischmann, B., Gietz, M. and Gnutzmann, S. (2004) 'Time-Varying Travel Times in Vehicle Routing', Transportation Science, 2: 160-173.

[4] Suh, W., Anderson, J., Guin, A., Hunter, M. (2015) 'Evaluation of Video Detection System as a Traffic Data Collection Method', Scientia Iranica 22 (6), S. 2092-2102

[5] Zhang G, Avery RP, Wang Y. (2007) 'Video-Based Vehicle Detection and Classification System for Real-Time Traffic Data Collection Using Uncalibrated Video Cameras', Transportation Research Record.;1993(1):138-147. doi:10.3141/1993-19

[6] Li, S., Yu. H., Zhang, J., Yang, K., Bin, R. (2014) 'Video-based traffic data collection system for multiple vehicle types', IET Intelligent Transport Systems 8(2):164-174

[7] Hasegawa, O., Kanade, T. (2005) 'Type classification, color estimation, and specific target detection of moving targets on public streets', Mach. Vis. Appl.,16, pp. 116–121

[8] Mallikarjuna, C., Phanindra, A., Ramachandra Rao, K. (2009) 'Traffic data collection under mixed traffic conditions using video image processing',Transp. Res. Rec. J. Transp. Eng.,135, (4), pp. 174–182

[9] Khan, S.M., Cheng, H., Matthies, D., Sawhney, H. (2010) '3D model-based vehicle classification in aerial imagery', Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1681–1687

[10] Tzutalin (2015) LabelImg Free Software: MIT License.

[11] Redmon, J., Farhadi, A (2015) 'You only look once: Unified, real-time object detection', arXiv preprintarXiv:1506.02640.

[12] Redmon, J., Farhadi, A. (2018) 'YOLOv3: An Incremental Improvement', cite arxiv:1804.02767Comment: Tech Report.

[13] Simonyan, K., Zisserman, A. (2015) 'Very Deep Convolutional Networks for Large-Scale Image Recognition', CoRR, abs/1409.1556.

[14] Zhou, X., Yao, C., Wen, H., Wang, Y., Zhou, S., He, W., Liang, J. (2017) 'EAST: An Efficient and Accurate Scene Text Detector'

[15] Baek, J., Kim, G., Lee, J., Park, S., Han, D., Yun, S., Oh, S. Lee, H., (2019) 'What Is Wrong with Scene Text Recognition Model Comparisons? Dataset and Model Analysis', International Conference on Computer Vision (ICCV)

Biography

Meng Jin (*1994) holds a Master`s Degree (M.Sc.) in Software Engineering from the University of Stuttgart with majors in computer vision. Since 2018 she has been working as research associate in Research and Innovation Center for Cognitive Service System (KODIS) of the Fraunhofer Institute for Industrial Engineering IAO.

Lars Andreas Mauch (*1987) has been a research associate at the Fraunhofer Institute for Industrial Engineering IAO in Stuttgart since 2017. As a member of the team "Energy Innovation" he focuses on city logistics and the electrification of urban commercial transport. Part of his research is the integration of decentralized and electrified fleets in transport systems. Before joining Fraunhofer IAO, he studied civil engineering at Karlsruhe University of Applied Sciences, where he worked for two years as a research assistant in the field of automated driving and machine-readable infrastructure.

Bernd Bienzeisler (*1964) holds a doctoral degree (Dr. rer. oec.) in business administration and is the head of the Research and Innovation Center for Cognitive Service System (KODIS) of the Fraunhofer Institute for Industrial Engineering IAO. His research focusses on data driven services and business models in different branches and industries.

2nd Conference on Production Systems and Logistics

Combining Process Mining And Simulation In Production Planning

Adrian Langer[1], Christian Ortmeier[1], Niels L. Martin[1,2], Tim Abraham[1], Christoph Herrmann[1]

[1]Institute of Machine Tools and Production Technology / Chair of Sustainable Manufacturing and Life Cycle Engineering / Technische Universität Braunschweig, Langer Kamp 19b, 38106 Braunschweig, Germany
[2]Center of Pharmaceutical Engineering (PVZ), Technische Universität Braunschweig, Franz-Liszt-Str. 35A, 38106 Braunschweig, Germany

Abstract

The conditions for industrial companies are changing due to increasing customer demands for individualised as well as sustainable products. Furthermore, companies are confronted with technological change by digital transformation. Therefore, production planning has to address various structural, procedural and organisational changes. Planning projects are often characterised by a high degree of complexity. In order to master the associated challenges, simulation models are used in production planning. In contrast to mathematical-analytical methods, simulation models examine and assess especially complex production systems and support improvement measures. A major difficulty during the model initialisation and the determination of the planning variables is the capture of data and the assurance of sufficient data quality. Both are associated with a high expenditure of time. At this point, manufacturing companies are faced with a conflict of objectives between the reduction of the planning time and the development of reliable simulation models. Process Mining (PM) can be used to capture data from central information systems and to uncover social and organisational networks and map them in a process model. This can create a well-founded data basis for simulation models.
To support simulation models within the planning process, a methodology linking process mining and simulation has been developed. This methodology improves the database within the planning process and renders it usable for rescheduling production systems. Potentials that can be achieved in the areas of data acquisition, data quality and model building are systematically analysed. The approach is validated on the basis of a use case from the pharmaceutical industry.

Keywords

Production Planning; Simulation; Process Mining; Data Acquisition

1. Introduction

Nowadays, production companies are confronted with decreasing batch sizes, highly fluctuating production rates, shorter product life cycles, shorter delivery times and an increasing number of product variants [1], [2]. To face these challenges, constant adjustments are necessary within the production system. This leads to an increasing effort in factory planning. For this reason, various approaches support the phases of factory planning by digital methods and tools [3], [4]. Simulations have proven to be a particularly valuable tool for decision support in factory planning. However, the data acquisition is associated with a very high effort. This results from the necessary data quality that must be available for accurate simulations. In this context, process mining (PM) offers the possibility to generate a data basis for simulations by analysing process-specific data.

For this purpose, past-oriented PM uses feedback data from existing IT systems. [5]
To support the initialisation of simulation models within the planning process, an approach that combines process mining and simulation is presented. The question is whether this approach can enable better simulation results and whether the effort in data acquisition and preparation can be significantly reduced. Therefore, the approach is applied in a case study of the pharmaceutical industry.

2. Theoretical background

2.1 Simulation models for decision support in factory planning

Factory planning encompasses the organisational, process-related and structural design or adaptation of the production system. Several approaches can be found in the literature that are based on a systematic, goal-oriented and phase-specific planning procedure that build on each other and combine the different planning tasks (e.g. process planning, layout planning, programme planning, etc.) [6], [7]. Basically, those planning phases develop from a draft to detailed planning and from an ideal to a real state. Starting from the definition of the goal, the further planning phases become more and more detailed. Subsequent planning phases can already begin even though previous ones have not yet been completed. In general, a distinction can be made between different basic planning cases [4], [6-8].
For each planning case, the classification in the planning level must be clarified. Starting from a rather global view for the network and site level, the following levels add more and more details reaching up to focus on the granular, detailed workplace of a specific area of a factory. Individual planning projects can either have an impact on several levels or be anchored in individual levels. The area structure level is responsible for the new planning or rescheduling of production chains and units within the production system. Here, the linking or arrangement of production stations is achieved via corresponding flow systems [3], [4], [6], [8]. The demands on the interdependencies and data also increase with increasing level of detail of the respective planning phases. With a high level of data availability and quality, more precise planning results can be achieved and the (re)planning project can be improved.

For further decision support, simulation tools can be used within the factory planning process. In general, models are an abstracted representation of the real system under investigation. In a manufacturing context, models can be classified based on the purpose of use. Jockisch and Rosendahl as well as Page and Kreutzer distinguish between descriptive, explanatory, forecasting and decision models. Besides the descriptive models, all other models are used in simulation modelling. Forecasting and decision models are the most common models in manufacturing for decision support. For simulation purposes, forecasting models use explanatory models to predict future scenarios. Decision models, which include optimisation models, have the purpose to facilitate the determination of options for action. For this purpose, the findings from an explanatory or forecasting model are supplemented by a decision space within the framework of decision models [9], [10].
In the context of factory planning, simulation has an important role with regard to the design and evaluation within the planning, realisation and operation of production systems to determine the operational performance. Depending on the planning case, the basic relationships of the production system already exist (brownfield) or are set up completely from scratch (greenfield). In rough planning, models can be initialised that have a high degree of abstraction, i.e. a low level of detail. As the planning process progresses, the demands on the models and thus also on the data and the associated level of detail grow. The biggest obstacle in creating models for real manufacturing applications is the difficulty in obtaining the right data and the determination of interrelationships [11]. The aim of modelling and simulation is to transfer conclusions gained through performance measurements on the model to the real system. In this way, simulation can support decision-making processes, which can assist in the selection and evaluation of solutions in the production system, especially in the case of several system variants related to the processual, organisational and structural interrelationships [12-14].

In addition, simulation models can be differentiated with regard to the paradigms chosen. The most important paradigms in simulation modelling are: System Dynamics (SD), Dynamic Systems (DS), Discrete Event (DE) and Agent Based (AB). Technically, SD and DS mostly deal with dynamic system behaviour, while DE and AB mostly deal with discrete system behaviour [15], [16]. In figure 1, typical simulation paradigms for specific elements of a production system are shown in relation to their degree of abstraction. In the context of factory planning, the necessary main elements have usually a higher degree of uncertainty and therefore a higher abstraction. According to this, the main elements for factory planning are highlighted in figure 1.

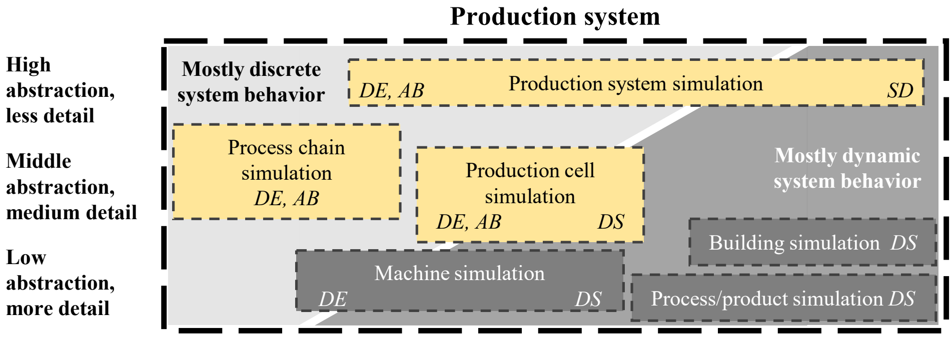


Figure 1: Overview of the different simulation models for different elements of a production system [16]

The definition of objectives in relation to the addressed planning tasks is particularly important in order to be able to select a suitable procedure for simulation. Material flows can be analysed via the discrete process chains by means of a DE or AB simulation, which changes the state only at certain events in a discrete manner. Ultimately, the approaches can also be combined depending on the planning task and objective. The combination of models for all elements of a production system, including product units, within a simulation would make it possible to analyse dependencies between the system elements involved and the effects of local improvement measures on the overall system [16]. In addition, a real-time simulation of the entire factory could run concurrently with real production operations and provide immediate results for short-term decisions [11]. However, a simulation is only as good as the input data. Process mining allows to access data of brownfield production systems.

2.2 Process mining

PM can be seen as a bridging technology between process science and data science. Process science includes, for example, techniques of workflow management, operations research and business process management. Within data science, besides techniques of statistics and data mining, also approaches of machine learning can be found. [17] PM techniques can be used to extract process-relevant information from an event log. These can generally be derived from Enterprise Resource Planning (ERP) or Manufacturing Execution Systems (MES). Event logs are a list of manual or system feedback data containing product and process specific information of a particular event. Van der Aalst et al. have defined five maturity levels for the data quality of an event log. Starting from the lowest maturity level, in which events only represent reality to a small extent and are increasingly paper-based, the highest maturity level records all events automatically and reliably. In addition to timeliness, the quality criterion of semantics is also met for all events [18].

By using algorithms, which are called miners in PM, a process model can be modelled based on the event logs. For this, an event log must contain at least a case ID, an event ID and a timestamp [18]. Furthermore, for a deeper evaluation of the process model, the description of the event (activity) as well as further order-, product- or process-specific information is useful. An example of event logs is shown in table 1.

Table 1: Event log of a data set

Case ID	**Event ID**	**Timestamp**	**Activity**	**Attributes**	**…**
Order number	Work station (WS)	Completion time	Process step	Quantity	…
010541528230	WS 01	14.12.2020 07:16	Preparation	7.000	…
010541528230	WS 02	15.12.2020 22:00	Production	7.000	…
010541528230	WS 03	15.12.2020 22:20	Post processing	7.000	…
010541528255	…	…	…	…	…

In process mining, van der Aalst distinguishes between the following types:

- *Process discovery* is used to create process models based on the event logs.
- *Process conformance* checks the degree of conformity between the process model and the actual observable process behaviour. The latter is described by the event logs.
- *Process enhancement* repeatedly links the created process model with event data to obtain further information, for example through performance analysis. In this way, errors in process models can be corrected or relevant decision points in the process can be focused [18].

Further, four different perspectives can be assumed in process mining, which can be the control-flow perspective, the organisational perspective, the case perspective and the time perspective [18]. Depending on the use case, different problems can be solved by using the types and perspectives of process mining. For the creation of a simulation model, for example, process discovery can be used to obtain a realistic model of the production processes. The purpose of this paper is to investigate the opportunities and obstacles of using process mining for simulations in the context of planning processes. For this, the current state of research is presented in the following.

3. Combination of PM and simulation

In literature, there are several methodologic approaches of combining process mining and simulation [5], [19], [20]. On the basis of the actual processes, Maruster and van Beest were able to determine the number of variants of the process flow as well as the average throughput times and incorporate these parameters into the simulation. The approach was applied in a Dutch government institution responsible for collecting fines and in the gas processing industry. In addition to this approach, the work of Aguirre et al. and Abohamad et al. also focuses on ensuring process transparency through PM. Thus, in both use cases, the actual process models could be described and averaged throughput times could be determined. The use cases were characterised by administrative tasks and activities. Aguirre et al. applied the developed process model to a procurement process in a university. Abohamad et al. used event logs in the healthcare sector for process analysis [5], [21]. The mentioned approaches are located in business process management and have delivered convincing results in this research field. Another approach by van der Aalst discusses the link between process mining and simulations and addresses the challenges faced by discovery techniques. Quality concepts such as recall, precision, and generalisation are focused here [22].
In contrast to these approaches, the focus of this work is on using PM for simulations to analyse an existing production system. The main difference is that target values from the work plan, such as processing and setup times, are already available for all process steps. In the manufacturing context, it can therefore be assumed that, due to a previous time recording, the planned processing times have a correspondingly high agreement with the actual processing times. By adding further influencing factors (e.g. machine failures, maintenance times, missing material), the results of the simulation should represent a well-founded image of reality. In the context of this work, it is therefore investigated whether improved simulation results can be achieved with data acquisition by PM.

4. Approach to combine process mining with simulations

The basic phases of a factory planning process are preparation, rough planning, detailed planning and implementation planning and execution. In the first phase, the goals and scope of the (re)planning project are derived. The actual simulation models can be used for rough and detailed planning to model the production system in order to gain knowledge about the production processes. Based on the results of the simulation, the final step is the implementation of the (re)planning project. The more accurately the results of a simulation match reality, the better an informed decision can be determined in the planning process. However, the more accurate the results, the more time and effort is required to collect and prepare the data. The simulation steps from rough to detailed planning are characterised for the planners by the comparison of effort and benefit. In practice, therefore, the effort is usually kept low and a simulation is generated on the basis of expert knowledge or planned values. Here, the approach aims at combining process mining with simulation, see figure 2.

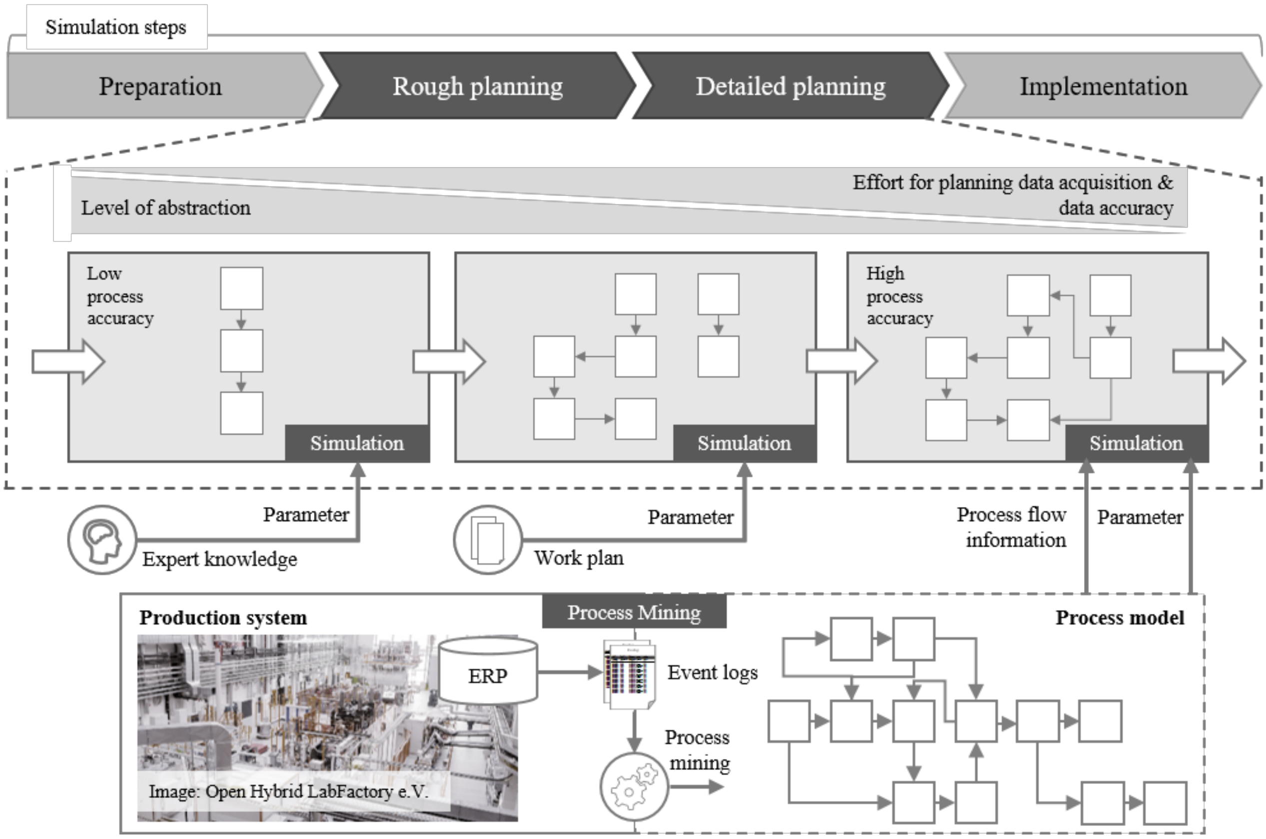


Figure 2: Approach to combine process mining with simulation

As shown in the approach, the PM can be used to create a process model based on the feedback data from the production system. In this context, the process model shows the actual production flow of manufacturing. This requires a correspondingly high level of data quality, and ensuring this is the core task of production data acquisition. Furthermore, parameters such as processing times, setup times, waiting times and failure times can be determined and used for the data basis. After an initial implementation of PM, these parameters can be continuously used for simulations with low effort and therefore, change the effort abstraction ratio for processes with high process accuracy. Furthermore, these parameters can be determined on a product-specific basis. In practice, product families are often formed in order to keep the effort low. Thus, in addition to reducing the effort, the data accuracy can also be increased. Hereby, it becomes clear that the support in the increasing data accuracy for the reduction of the associated efforts in the acquisition has a great impact on the planning processes. Ultimately, these conditions allow saving time as well as costs.

5. Use case

For this use case, real ERP order data for the final packaging of pharmaceutical products were received and will not be described in detail due to non-disclosure agreements. The data was cleared up of doublets as well as faulty datasets and finally comprises 912 orders for a whole production year (first half: 443; second half: 469) for four packaging lines. Besides the order specifics, the order data includes also the individually planned and real production time, such as process interruptions for each order produced on several process lines.

As illustrated in figure 3, the data is split into the first and second half of the year in order to allow the scenario of having the first half of the years' real production data and the planned data for the second half of the year. The process mining approach is applied on the production data and allows the reference with the second years' production data, which will be compared with the simulation data resulting from the input of the process mining data as well as the planned data. The comparison allows to conclude to what extent the planned data differs from the real production data compared to the use of PM.

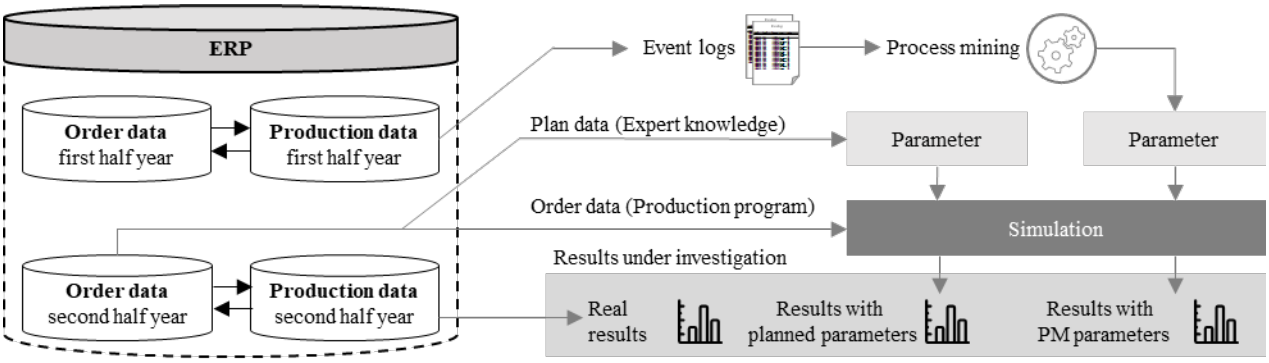


Figure 3: Comparison of real production data versus simulation scenarios by planned data and PM

A schematic overview of the agents and their interactions is illustrated in figure 4. The simulation is modelled in AnyLogic© using an agent-based simulation approach. The modelled agents of this model are a bulk storage for the raw, solid pharmaceutical products, the different process lines and an order scheduling. The bulk storage is able to dispatch products to the process lines and replenish the storage by defined amounts depending on the pharmaceutical product if an order from the process lines extends the stock of this product. The order scheduling assigns incoming orders to the process lines according to the lines' feasibility of handling these orders and the earliest possible production start at the order incoming date.

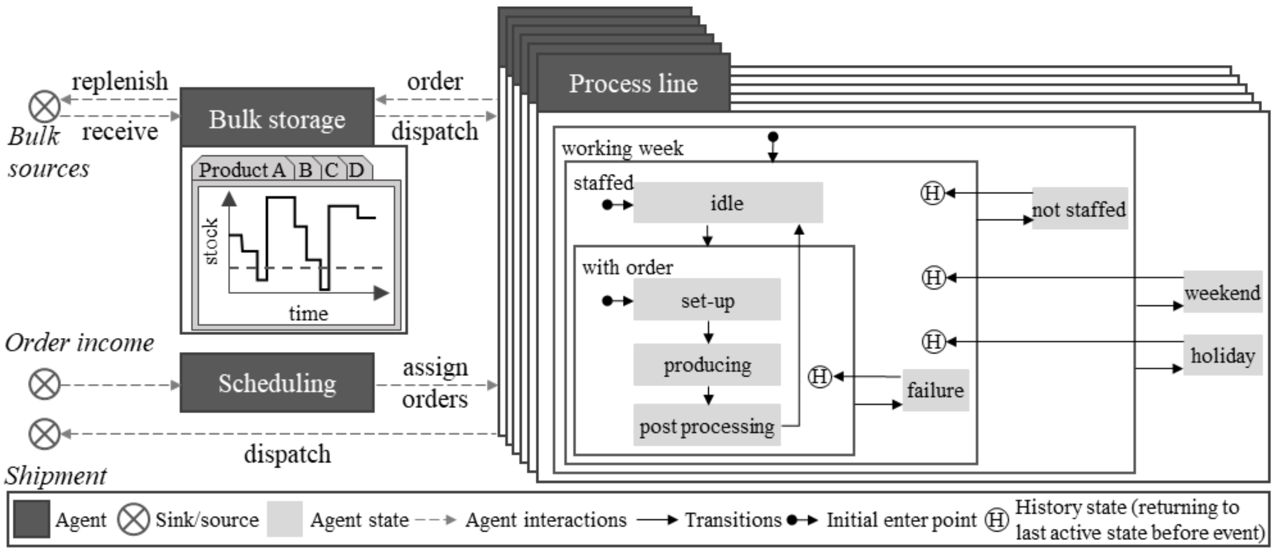


Figure 4: Schematic overview of the simulation model for the use case

The process lines consider weekends and public holidays and shift times to have a realistic and similar setting compared with the real production data on a macro scale. Those settings are the same in the three scenarios. On the micro scale, the process lines are in idle state until the order queue of the process line is filled and the order will be processed. If a process line processes an order, four states are possible: set-up, production, post processing and failure. For each state, there are parameters describing these states. Those parameters differ depending on the PM and the planning data and are listed in table 2.

Table 2: Applied process mining and planned data for simulation scenarios

State	Parameter	PM data [order data, 1. half]	Planned data [order data, 2. half]
Set-up	Set up time	Average set-up time of a specific process line	Planned set-up time of an order from order data
Production	Production ratio	The assessed production ratio of former orders per product group and line	The order's planned production rate
Post-processing	Post processing time	Average post processing time of a specific process line	Planned post processing time of an order
Failure	Failure times	Are included in production ratios	Are not considered in planning

The production ratio used in the PM and the planned data scenario are compared with the real order production ratio, which is shown in figure 5. The orders are sorted by rising planned production time. It can be observed that the real time data sets have outliers (A), which occur when the data logs are not logged properly for diverse reasons. Even though proper data sets are desirable, in practice, they are often hard to achieve and in this study they are not excluded. Furthermore, the production rates come closer to the real processes through data acquisition with the help of PM than can be achieved with the planning data. It can be observed that the real production rates for small and medium planned production rates are mainly too low, while the planned production rates for higher production rates match the real production rates better. This effect cannot be observed for the comparison of PM with the real production rates. Here, the real production rates seem to oscillate around the PM production rates with minor discrepancies.

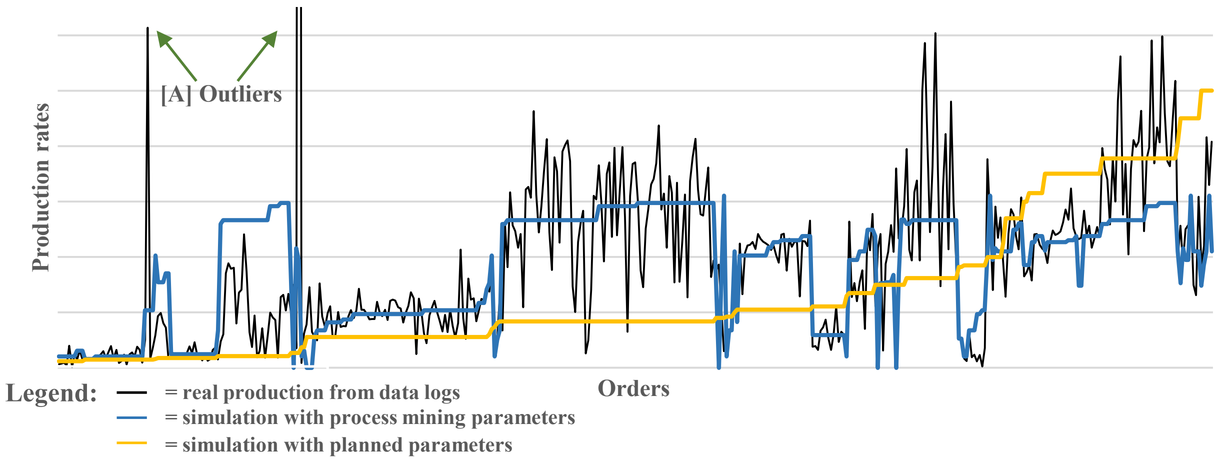


Figure 5: Comparison of PM and planned production ratios with the orders' real production ratio

This already shows that the process mining approach allows a better approximation of the input parameters for the simulation study. The diagrams in figure 6 show the deviations of the production volume and the throughput time from the simulation in relation to the real data. Regarding the above mentioned production ratio, the production volume shows a similar outcome. The real data sets and the data sets achieved by PM can be presented with a high amount of agreement. Here, however, the planned values are again found with

a higher deviation (figure 6, a). The middle area is particularly striking here; without offsetting the outliers, a very high agreement of the averaged production rates via PM compared to reality can be achieved. In comparison, the planned values are constantly below the real values. While for all three results the production volume is the same, it can be observed that the time needed for the completion of all orders is more accurate for the simulation results with PM in comparison to the planned data input. In this case, this leads to an overestimation of the required production time.
The diagram on the right side shows the cumulative deviation over time following the orders' chronological completion for the throughput time (figure 6, b). Here, it is particularly clear that the deviation increases as the number of orders increases for the simulation results using real and planned data. However, the deviation of the simulation results using real and PM data also exists, but the deviations compensate each other better. This effect is most probably a result of the usage of average production rates for product groups on specific production lines. Using a more specific approach might lead to even better results and should be a part of future research. Furthermore, these graphs allow the statement that the simulation itself is capable of forecasting the close to real throughput times using PM parameters.

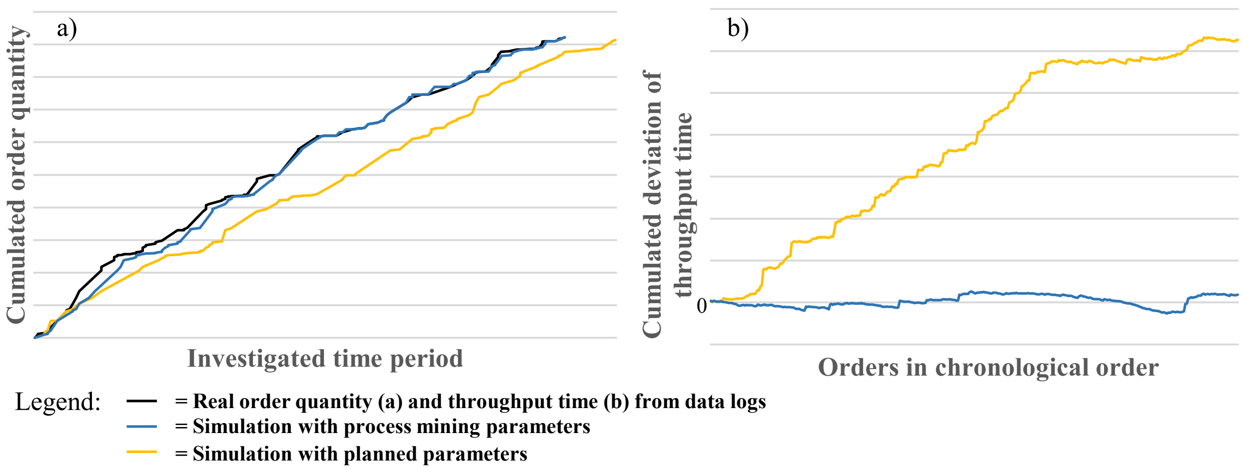


Figure 6: Comparison of real with PM and planned data for (a) the product volume and (b) the throughput time

Overall, this shows how strongly the planned values in this use case deviate from the real values. With PM, the parameters that serve as input for the simulation can be determined and transferred much more precisely. In the context of production planning, this is advantageous for the process-related, organisational and structural design.

6. Conclusion and outlook

In this work, the benefits of PM for ensuring a solid data basis for simulations were demonstrated. In an extensive use case of the pharmaceutical industry, process mining could be used for data acquisition and processing on the basis of an existing production line. The results demonstrated that, even in production planning, data is often not sufficiently accurate to develop a simulation and derive decisions based on it. In addition, the effort in data provision and procurement can be reduced by PM. Especially after the initial implementation of PM, the simulation can be adjusted continuously and with little effort. Due to the high prediction accuracy of processing and throughput times, new potentials also arise for the subject area of production planning and control. Here, for example, PM can be used in production control for targeted order release in order to increase adherence to schedules. In addition to rescheduling, data acquisition with PM can be used as a data basis for new facility planning projects within factory planning. Based on the process-related throughput times, production rates and downtimes, a more detailed target description including the scope of functions can be designed.

Acknowledgement

This paper evolved of the research project “Synus” (Methods and Tools for the synergetic conception and evaluation of Industry 4.0-solutions) which is funded by the European Regional Development Fund (EFRE | ZW 6-85012454) and managed by the Project Management Agency NBank.

References

[1] Gaal, A., Hofer, K., Ryback, T., Lingitz, L., Sihn, W., 2020. An agent-based approach to planning and controlling adaptable cell-oriented assembly systems. Procedia CIRP 93, 1158–1163.

[2] Tolio, T., Ceglarek, D., ElMaraghy, H.A., Fischer, A., Hu, S.J., Laperrière, L., Newman, S.T., Váncza, J., 2010. SPECIES—Co-evolution of products, processes and production systems. CIRP Annals 59 (2), 672–693.

[3] Pawellek, G., 2008. Ganzheitliche Fabrikplanung: Grundlagen, Vorgehensweise, EDV-Unterstützung. Springer-Verlag, Berlin, Heidelberg.

[4] Grundig, C.-G., 2018. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 6., neu bearbeitete Auflage ed. Hanser, München, 333 pp.

[5] Aguirre, S., Parra, C., Alvarado, J., 2013. Combination of Process Mining and Simulation Techniques for Business Process Redesign: A Methodological Approach, in: Cudré-Mauroux, P., Ceravolo, P., Gašević, D. (Eds.), Data-driven process discovery and analysis. International Symposium, SIMPDA 2012, Campione d’Italia, Italy, June 18-20, 2012, vol. 162. Springer, Berlin, pp. 24–43.

[6] Bauernhansl, T., 2020. Fabrikbetriebslehre 1: Management in der Produktion. Springer Vieweg, Berlin, 388 pp.

[7] Verein Deutscher Ingenieure, 2011. VDI 5200 Part 1: Factory planning - Planning procedures. Beuth Verlag GmbH.

[8] Schenk, M., Wirth, S., Müller, E., 2014. Fabrikplanung und Fabrikbetrieb: Methoden für die wandlungsfähige, vernetzte und ressourceneffiziente Fabrik, 2., vollständig überarbeitete und erweiterte Auflage ed. Springer-Vieweg, Berlin, Heidelberg, 832 pp.

[9] Jockisch, M., Rosendahl, J., 2010. Klassifikation von Modellen, in: Bandow, G., Holzmüller, H.H. (Eds.), „Das ist gar kein Modell!“. Unterschiedliche Modelle und Modellierungen in Betriebswirtschaftslehre und Ingenieurwissenschaften. Gabler Verlag / Springer Fachmedien Wiesbaden GmbH Wiesbaden, Wiesbaden, pp. 23–52.

[10] Page, B., Kreutzer, W., Gehlsen, B., 2005. The Java simulation handbook: Simulating discrete event systems with UML and Java. Shaker, Aachen, 502 pp.

[11] Fowler, J.W., Rose, O., 2004. Grand Challenges in Modeling and Simulation of Complex Manufacturing Systems. SIMULATION 80 (9), 469–476.

[12] Negahban, A., Smith, J.S., 2014. Simulation for manufacturing system design and operation: Literature review and analysis. Journal of Manufacturing Systems 33 (2), 241–261.

[13] März, L., Krug, W., Rose, O., Weigert, G., 2011. Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen. Springer, Heidelberg, 220 pp.

[14] Verein Deutscher Ingenieure, 2014. VDI 3633 Part 1: Simulation of systems in materials handling, logistics and production - Fundamentals. Beuth Verlag GmbH.

[15] Borshchev, A., Filippov, A., 2004. From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools. The 22nd International Conference of the System Dynamics Society.

[16] Schönemann, M., Bockholt, H., Thiede, S., Kwade, A., Herrmann, C., 2019. Multiscale simulation approach for production systems. Int J Adv Manuf Technol 102 (5-8), 1373–1390.

[17] van der Aalst, W., Damiani, E., 2015. Processes Meet Big Data: Connecting Data Science with Process Science. IEEE Trans. Serv. Comput. 8 (6), 810–819.

[18] van der Aalst, W., Adriansyah, A., Medeiros, A.K.A. de et al., 2012. Process Mining Manifesto, in: Daniel, F. (Ed.), Business Process Management Workshops. BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011; revised selected papers, vol. 99. Springer, Berlin, pp. 169–194.

[19] Măruşter, L., van Beest, N.R.T.P., 2009. Redesigning business processes: a methodology based on simulation and process mining techniques. Knowl Inf Syst 21 (3), 267–297.

[20] Rozinat, A., Mans, R.S., Song, M., van der Aalst, W., 2009. Discovering simulation models. Information Systems 34 (3), 305–327.

[21] Abohamad, W., Ramy, A., Arisha, A., 2017. A hybrid process-mining approach for simulation modeling, in: Chan, W.K., D'Ambrogio, A., Zacharewicz, G., Mustafee, N., Page, E.H., Wainer, G. (Eds.), WSC'17. Proceedings of the 2017 Winter Simulation Conference, December 3-6, 2017, Las Vegas, NV, vol. 117. IEEE; Omnipress, Piscataway, NJ, Madison, WI, pp. 1–12.

[22] van der Aalst, W., 2018. Process Mining and Simulation: A Match Made in Heaven! Proceedings of the 50th Computer Simulation Conference 50, 1–12.

Biography

Adrian Langer (*1993) studied industrial engineering with a specialisation in mechanical engineering at the TU Braunschweig. Since 2019, he has been working as a research assistant at TU Braunschweig in the sustainable production and factory systems working group of the Institute of Machine Tools and Production Technology (IWF) with a special focus on simulation of production systems.

Christian Ortmeier (*1992) studied industrial engineering with a specialisation in mechanical engineering at the TU Braunschweig. Since 2019, he has been working as a research assistant at TU Braunschweig in the sustainable production and factory systems working group of the Institute of Machine Tools and Production Technology (IWF) with a special focus on data-driven methods and tools in lean administration.

Niels Lasse Martin (*1988) studied industrial engineering with a specialisation in mechanical engineering at the TU Braunschweig. Since 2017, he has been working as a research assistant at TU Braunschweig in the sustainable production and factory systems working group of the Institute of Machine Tools and Production Technology (IWF) with a special focus on pharmaceutical production.

Dr. Tim Abraham (*1987) studied industrial engineering with a specialisation in mechanical engineering at the TU Braunschweig. Since 2014, he has been working at the IST, Fraunhofer Institute for Surface Engineering and Thin Films completing his PhD in 2019. Since 2020, he is head of the sustainable production and factory systems working group of the Institute of Machine Tools and Production Technology (IWF).

Prof. Dr.-Ing. Christoph Herrmann (*1970) is university professor for Sustainable Manufacturing & Life Cycle Engineering and co-director of the Institute of Machine Tools and Production Technology (IWF), Technische Universität Braunschweig as well as director of the Fraunhofer Institute for Surface Engineering and Thin Films IST since November 2018.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Production Planning And Control In Distributed And Networked Open Production Sites – An Integrative Literature Review

Lennart Hildebrandt[1], Tobias Redlich[1], Jens P. Wulfsberg[1]
[1]Helmut-Schmidt-University, Hamburg, Germany

Abstract

The COVID-19 pandemic has once again drastically highlighted the trend and need towards urban and distributed production in cities (so-called Fab Cities) and their importance for society in order to independently meet the demand for physical goods. For small but highly individualized products, the manufacturing process is now possible in distributed and open production sites (so called FabLabs) equipped with digital manufacturing machines. These places empower individuals, start-ups, SMEs or companies to innovate, produce and educate. However, many open production sites are operated independently of each other, reducing resource efficiency, capacity utilization and competitiveness. This strives against the trend of physical and digital networking, which the manufacturing industry has long since completed in order to use its capacities more efficiently. In this paper, an integrative literature review is used to hypothesize and verify that such production planning and control (PPC) for open, distributed and bottom-up controlled production networks has not yet been scientifically researched. As a result of the review, it appears that today's production can be divided into three main types. The first main type represents the closed factory with its own PPC. The second main type represents globally connected and distributed value networks (e.g., Industry 4.0, cloud manufacturing) that are controlled top-down. The third and largely unexplored main type consists of open, bottom-up controlled as well as locally distributed but globally connected open production sites. To increase the future competitiveness and resilience of a sustainable Fab City, the authors show that further research is needed on the controlling and governance of open and urban production sites which the authors present in a research agenda.

Keywords

Production planning and control; Production network; Cloud manufacturing; Decentralization; Fab City

1. Introduction

The COVID-19 pandemic has highlighted the fragility of the current production system in unprecedented situations of global shortages in the medical, protective equipment and other industries. This is the result of globalized value chains in which physical products are manufactured and then transported across the globe. This can be problematic: The medical and protective equipment sector was drastically curtailed in the Asian region in the spring of 2020 due to COVID-19, which resulted in supply shortages of face masks, respiratory equipment and face shields. [1,2] In response, individuals and makers began to develop protection products (e.g., personal protective equipment) as part of globally networked online communities, shared knowledge and subsequently manufactured them in open production sites (FabLabs or OpenLabs). FabLabs or

OpenLabs are production sites equipped with professional and digital manufacturing infrastructure (e.g., 3D printers, CNC mills) that are open to the public. As a result, such places basically offer the possibility of producing smaller everyday objects in a resilient manner on a small scale. [3-5] Some of the products manufactured there have even been officially approved and have thus made a real contribution to overcoming the pandemic. [6,7]

What could be exemplarily tested and successfully implemented during the pandemic is part of a new and global trend since 2011, which can be summarized under the term "Fab City". This concept comprises digital networking and the open exchange of product data (e.g., design files, bill of materials) and knowledge between cities (Fab Cities) within the network. As a result, all cities that are part of this network have access to the data and knowledge, which, in combination with publicly accessible, local and distributed production sites, enables them to physically manufacture a product. Cities are thus globally connected and locally productive. [8] Due to the topicality of the issue, public attention regarding this topic is increasing.

However, distributed manufacturing within a city is not as simple as might be imagined. Many FabLabs in Fab Cities are operated independently, which has a negative impact with regarding to resource efficiency and in competition with industrial value creation models. In order to make a relevant contribution to the supply of a city with physical products and to reduce the effects of global and ecologically unsustainable transportation chains, the goal should be to increase the efficiency and productivity of the individual open production sites by networking and controlling these decentralized production capacities analogously to the manufacturing industry. Through such control of the decentralized FabLabs, it would be possible to systematically capture demands from the city, pass it on to available makers and machines and produce locally. But how can production be planned and controlled in a city with such diverse actors, and what approaches already exist in this field? We hypothesize that such suitable, open and transparent production planning and control (PPC) systems that connect the needs of individuals, SMEs, startups and other actors with the local and distributed manufacturing capacities in the city do not currently exist.

In this paper, our objective is to examine this hypothesis. As a method, we use an integrative literature review. By doing so, we aim to determine the current state of literature on the topic. Due to the topicality of the subject in the field of decentralized and local production, there are currently many new publications. Therefore, we aim to analyze the state of the literature in the area of production planning and control in networks and check whether our hypothesis, that there is no PPC for Fab Cities yet, is true. In this paper, we first present the methodology and an overview of the literature. In the second step, we discuss the main results and, building on this, we present a research agenda in the third step.

2. Methodology

An integrative literature review includes and critiques previously published findings on a complex topic to gain a deeper understanding and new insights. [9,10]. We follow the approach of Torraco [10] and built our research in two steps: identifying and collecting papers via a keyword search in a database that are most representative regarding the topic, and then briefly defining and reviewing the keywords and analyzing the theoretical basis.

As a database we have used Web of Science. Web of Science offers up to 10 different sub-databases to choose from, which can be searched for specific terms. It is possible to filter by source type and discipline. In total, Web of Science lists over 21,100 journals, over 180,000 conference proceedings and over 80,000 books. The keywords were selected through a preliminary research. They represent the subject area of production planning and control, which primarily deals with individual factories, as well as the extended production planning and control in networks through the more general terms "Production Networks" and "Cloud Manufacturing".

In the course of the keyword search, a large selection of literature was found in the database, thus, the selection of results was further limited using filters. For this purpose, on the one hand, the exact and coherent word group of the individual keyword was searched for, and, on the other hand, the condition that the keyword had to be listed in the title or abstract of the respective paper was introduced. This significantly reduced the number of results. (see Figure 1)

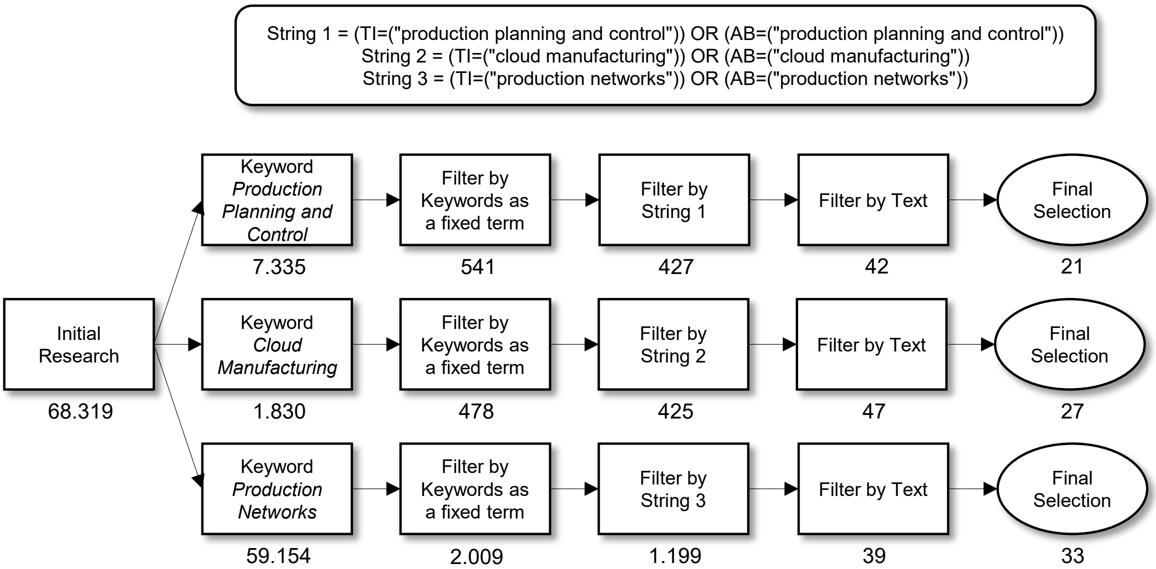


Figure 1: Literature selection process

In the further procedure of literature selection, the title and abstract of each result were considered as a first step. As a second step, the paper was read in full and, if relevant, the central statements and aspects were elaborated, and the sources in the bibliography and the quality of the journal were checked. In particular, the selection process discarded literature that had no overlap with the present work in the title or abstract. Furthermore, literature that was not peer-reviewed and/or controlled by a publisher was excluded.

In the following, the results of the literature research are presented. First, all three keywords are defined for this purpose based on the statements of the selected literature (chapter 3). This is followed by an analysis of the individual papers (chapter 4). Finally, the main findings (chapter 5) and the research agenda (chapter 6) are presented.

3. Keyword definition

3.1 Production Planning and Control (PPC)

In a manufacturing company, the PPC is responsible for planning and controlling a manufacturing process in terms of schedules, capacities and output. [11,12] The PPC converts a customer order into a physical product. [13] It is thus essential for meeting customer requirements regarding production parameters and can therefore be found in different areas of a company and in various designs. [14,15] Typical functions of a PPC include demand planning, demand control, capacity planning and order planning. The main objectives of these functions are to reduce work-in-process, reduce lead times, reduce inventory costs, improve responsiveness to changes in demand and meet delivery deadlines. [15] The systems are basically hierarchical and top-down. [13,16]

The use of PPC has changed a lot over the years. In the past, they were very static and only insufficiently considered changes from the current situation (real-time representation). [17-19] Today's factories, in contrast, require agile, scalable and reconfigurable production systems. [11,20] For this purpose, the integration of the Internet of Things (IoT) and further sensor technology is promising. [16,21] Furthermore, since about 2005, production has increasingly been "make-to-order", meaning a product (or a part of it) is not manufactured until the customer order has been received, so that essential product features can still be adapted to customer needs. [22]

3.2 Cloud Manufacturing (CMfg)

Conventional PPC is not sufficient for distributed production. Due to globalized value chains, extensive networking in production is necessary which can be achieved by CMfg. [23] Although there is no unified definition for this term, it basically refers to the principle of cloud computing in the manufacturing world. [24,25] CMfg is a service- and demand-oriented production approach in networks, where configurable production resources and capacities can be accessed in real time via the Internet. [26-28] This means that production resources and capacities are virtualized and can thus be accessed by any user. [29] The concept aims to circulate capacities and to realize demand-driven utilization of various production resources and capacities. This is done by providing reliable, high-quality, cost-effective and demand-driven manufacturing services for the entire manufacturing life cycle. [27,30]

Production can thus take place in a distributed manner, while all the necessary services and knowledge (e.g., production data, scheduling, business workflow management) of the network are stored centrally in a cloud. Additionally, any organization or individual should be able to participate in the network and share knowledge in the cloud, similar to Wikipedia. [24,27] The technical implementation is supported in the physical and virtual domains by IoT, cloud computing and other computing technologies. [27,31]

There are three different roles in CMfg. The provider offers and owns the production capacities in the respective production. The operator runs the platform and thus provides services in the cloud for all users. The consumers are the subscribers of the platform who buy manufacturing cloud services from the operator. [27] Whether anyone (e.g., individuals or OEMs) or only business partners can be consumers or providers is handled differently. [32,33] However, unlike Industrie 4.0, CMfg is fundamentally not seen as a purely industrial concept but is usually implemented that way. [34,35]

3.3 Production Networks (PN)

Production Networks (PN) or Global Production Networks (GPN) are organized networks in an area, consisting of a company and non-company institutions, through which goods and services are produced and distributed. [36,37] For the joint production of these goods and services, all activities in the network are combined and shared. [38]

The network benefits from this division because the modern economy consists of networked and specialized production units that share knowledge and with complementary capabilities in each unit allow the whole network to cover a wider range of capabilities. [39,40] As a result, production networks provide rapid and cost-effective access to resources, skills and knowledge, reducing transaction costs for each network partner and enabling the network as a whole to respond more quickly to market trends. [40-43]

Production networks are very dynamic and variable due to constant external changes [36], but at the same time they are very stable in the long term and have a polycentric hierarchical structure. [44] Unlike global commodity chains and global value chains, production networks do not have a linear structure. Instead, all possible network configurations between network partners are allowed, making the classical customer-supplier relationship obsolete. [36,44] Companies are the main active partners in production networks, but theoretically all relevant actors and relationships are addressed, which, however, is usually not the case in practice. [36,40,45]

4. Literature review

After defining the three keywords according to the selected literature, the analysis of the literature follows below. A total of 81 papers are evaluated in this review. The chronological sequence of publications clearly shows that the majority of publications was published after 2010. This is due to the novelty of the concepts, which were first made possible by technological progress (e.g., cloud computing, IoT, ICT). The classic PPC has been around for a longer time and is well established in the manufacturing sector. (see Figure 2)

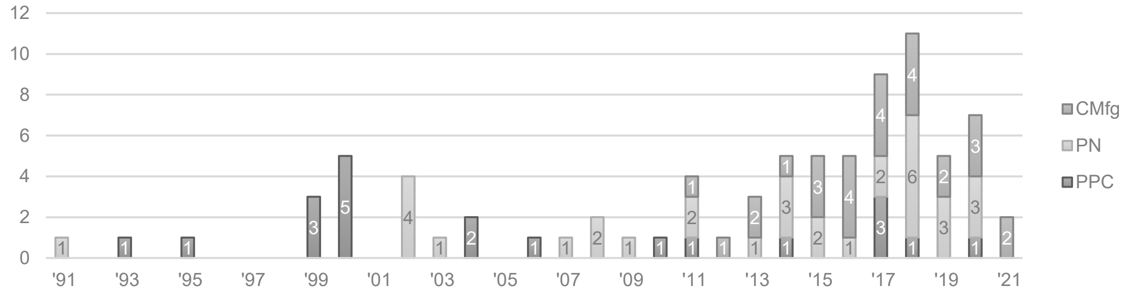


Figure 2: Chronological representation of the publications for the individual keywords

The main statements were extracted from the selected literature and evaluated in terms of accessibility, networking and decentralization. It was found that although many authors from the various approaches describe or strive for (global) networking, these networks are mostly closed or limited to a specific group of users. In particular, economic networks within companies and industries were described. Globally networked production systems with accessibility for everyone (individuals, SMEs, startups, companies, ...) could only be recognized in rudimentary form in exceptional cases. (see Figure 3)

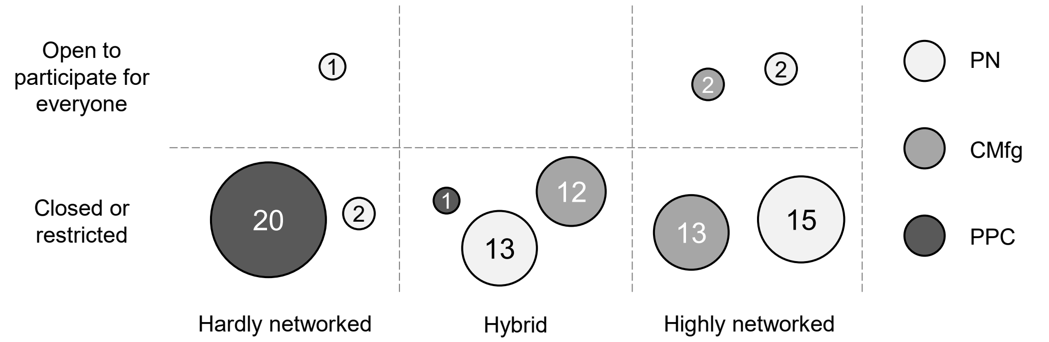


Figure 3: Categorization of papers according to openness and connectedness

The situation is similar regarding the decentralization of the individual sites (both global and local). Here, too, it can be seen that the majority of authors describe or strive for decentralization of the sites, but opening up this production network to outsiders is also not described or strived for in the vast majority of cases. (see Figure 4)

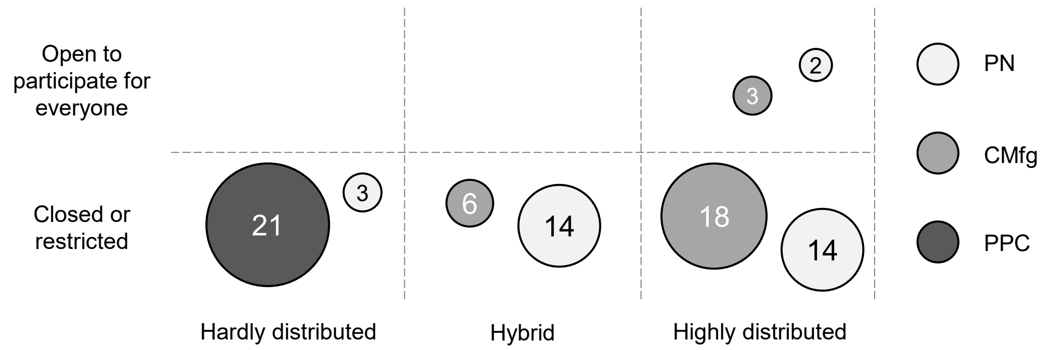


Figure 4: Categorization of papers according to openness and distributivity

The papers were published in various journals covering a range of different disciplines. However, the majority is focused on the field of production engineering. Since the topic is fundamentally interdisciplinary, additional disciplines could also be identified. The following table shows all journals including the editors who published at least two papers mentioned in this review. (see Table 1)

Table 1: Journals with more than one paper in this review

Journal	Total counts	Publisher	Discipline	Impact Factor
Production Planning & Control	8	Taylor & Francis	Operations management	4.22
The International Journal of Advanced Manufacturing Technology	6	Springer	Manufacturing Technology	2.96
International Journal of Production Research	6	Taylor & Francis	Production Engineering	4.58
International Journal of Computer Integrated Manufacturing	5	Taylor & Francis	Engineering	2.48
Journal of Economic Geography	3	Oxford Academic	Economic geography	3.29
Journal of Manufacturing Science and Engineering	3	ASME	Manufacturing Technology	2.88
CIRP Annals	2	Elsevier	Manufacturing Technology	5.52
IEEE Access	2	IEEE	Multidisciplinary	4.64
CIRP Journal of Manufacturing Science and Technology	2	Elsevier	Manufacturing Technology	3.26
Regional Studies	2	Taylor & Francis	Urban and regional change	3.34
Research Policy	2	Elsevier	Management and Technology Science	5.35

Most of the contributions present an overview and basics on the respective topic. Since PPC has already been an established tool in production for years, the authors here deal in particular with the simulation, modeling and flexibilization of production and the presentation of new concepts. Regarding cloud manufacturing, the main focus is on presenting the basics of the concept. Since the concept is still quite new (first mentioned in 2009), mainly theoretical approaches are described. Not much is published on the implementation or analysis of existing networks. This is in contrast to production networks, for which there is a detailed analysis of practical implementation and realization. Apart from this, the basics are also presented here. There are only a few publications on the integration of customers in the product development process and on production in maker networks. And although the topic of sustainability has become ubiquitous in recent years, only two authors of the selected papers analyzed effects of production networks on ecology. (see Table 2)

Table 2: Main categories of papers by keywords

Keyword	Topic	Reference
Production Planning and Control	Flexibilization	[13,20,46,47]
	Increase of efficiency or production	[48]
	Real-time production	[49,50]
	Framework	[51]
	Modeling and simulation	[14,17,19,22,52,]
	Conceptualization	[16,53–57]
	Implementation	[58,59]
Cloud Manufacturing	Basics and overview	[25,27,29,31,35,60.61]
	Modeling and Simulation	[26,62,63]
	Optimization and increase of efficiency	[64–67]
	Implementation	[68]
	Future perspectives	[23,24,33,34]
	Real-time production	[69]
	Dynamization	[28]
	Surveys and studies	[70,71]
	Collaborations	[72]
Production Networks	Analysis of networks	[45,73–78]
	Influences on networks	[79,80]
	Surveys and studies	[81,82]
	Increase of efficiency or production	[83]
	Customer integration	[84,85]
	Basics and overview	[36,38,39,44,86-93]
	Production Planning and Control	[94]
	Sustainability	[95,96]
	Flexibilization	[97]
	Knowledge Management in Networks	[40,98]

5. Results

Based on the definitions and the analysis, we identified essentially three relevant main types of production, which differ massively in terms of their accessibility, distribution and degree of connectedness. Moreover, the actors in the main types are very different.

Main type 1 represents **closed production processes** within a company which are hardly or not networked and distributed. These are production companies that produce at one location and are controlled by a central PPC. In these companies, it is top-down specified which production step has to be carried out at which point in time, by whom and in which sequence. There is no provision for opening up the production to external parties. Examples of such companies are traditional manufacturing SMEs in Germany. There are many publications about the main type 1.

Main type 2 represents (largely) **closed but highly networked and distributed** companies with several production sites or a network of several companies with distributed production sites. In these primarily commercially oriented networks, a central body dictates which production step is to be carried out when, where and by whom. Opening up to outsiders is largely only intended if they can improve the efficiency of the network through additional production capacities. This primarily addresses other companies. Opening up to individuals is generally not intended. In addition, decisions in these networks are also top-down driven. Producers (companies and employees) are not involved in the decision-making process. Due to the great distances between locations in globally distributed production, additional logistical effort is required. Examples of such companies can be found in the traditional supplier industry (e.g., automotive and aircraft industry), where large corporations (e.g., VW, Airbus) stipulate the specifications for the production centrally and have them implemented globally. There is an increasing amount of literature on this main type.

Main type 3 represents **openly accessible, networked and distributed production** opportunities and sites that are accessible to everyone and address different actors (e.g., individuals, SMEs, startups, companies) in terms of production, innovation and education due to the number and professionalism of the machines. [3,82] These production sites (FabLabs or OpenLabs) and smaller factories can be embedded in the urban environment due to their small size as well as low noise pollution and emission levels, allowing cities to produce again at the point of need. Networks with open production sites offer society the opportunity to participate in and co-determine value creation again. [99-101]

6. Research Agenda

As this review shows and as it was hypothesized in the introduction, there are currently very few publications on production planning and control in open, distributed and networked production sites that are simultaneously controlled by the entire network (bottom-up). Based on the social, economic and environmental potentials that are already described in the literature in the areas of production, innovation and education (e.g., applied STEM education for youth, commercial use by artisans to individualize products, prototyping by SMEs and startups, manufacturing of PPE during COVID-19 pandemic), we see a need for further research on production planning and control of open, networked, and distributed production sites in order to promote this specific main type and build a new resilient and competitive production system. Specifically, we recommend the following three sequential research steps:

1. **Analysis of the current state** in open production sites with regard to the machines and social networks as well as the local and global networking of the individual sites. To learn more about this current state, we propose an analysis of the states in different Fab Cities. The individual Fab Cities can be seen as contexts and cases in which new insights can be created by means of interviews, participant observation and existing production data, which can serve as the basis for an urban production planning and control system (step 2).
2. **Derivation of research work** in terms of modeling and simulation for the planning and control of such an open and dynamic production network. Modeling and simulation are carried out on the basis of the knowledge gained in the first step. This enables analyses in terms of increasing efficiency, productivity and profitability, which underline the added values for the urban stakeholders through an urban and decentralized production planning and controlling system.
3. **Exemplary implementation** of such a system in a production network with subsequent field tests to research the practical suitability of such a production system. The implementation is carried out in the analyzed cities, providing a match between simulation and reality, so that the practical benefits can be evaluated and confirmed by a PPC in open and distributed production networks.

7. Conclusion

The COVID-19 pandemic has shown how important resilient and locally embedded but globally connected production sites are in order to share data globally and produce products locally. But it has also been shown that the efficiency and penetration of the networks have been mediocre because the joint activities have been poorly planned and managed. We therefore hypothesized that there is still no suitable production planning and control system for such production to be found in the body of literature.

To test this hypothesis, we conducted an integrative literature review including 81 papers. We found that there are three main types of production forms under the keywords "Production Planning and Control", "Cloud Manufacturing" and "Production Networks", which differ significantly in terms of their openness, distribution and networking.

- **Main type 1** represents closed and local production within a company.
- **Main type 2** represents closed and economic networks with local and global distribution and connections, which cannot be used by outsiders.
- **Main type 3** represents an open, distributed and networked production system for everyone.

Main type 3 addresses the needs of all actors (individuals, SMEs, startups, companies) through the number and professionalism of machines. In this respect, there is very little literature on the planning and controlling of such a network. Therefore, we can confirm the hypothesis through the integrative literature review.

We see a need for further research in the area of production planning and control in open, distributed and networked production sites to promote the development of a resilient, new and competitive manufacturing field. Specifically, we recommend that researchers in the field of production engineering survey the current state in such networks in order to model and simulate a PPC system and then implement it in a network as an example and study it within the framework of a field test.

References

[1] Buckley, C., Sui-Lee, W., Qin, A., 2020. China's Doctors, Fighting the Coronavirus, Beg for Masks. nytimes.com.

[2] Pearce, J.M., 2020. A review of open source ventilators for COVID-19 and future pandemics. F1000Research 9, 218.

[3] Hildebrandt, L., Moritz, M., Seidel, B., Redlich, T., Wulfsberg, J.P., 2020. Urbane Mikrofabriken für die hybride Produktion. ZWF 115 (4), 191–195.

[4] 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. 05.01.2016 - 08.01.2016. IEEE, pp. 707–715.

[5] Redlich, T., Moritz, M., Buxbaum-Conradi, S., Krenz, P., Heubischl, S., Basmer-Birkenfeld, S.V., 2015. OpenLabs - Collaborative Industrialization with Distributed and Open Source Microfactories. AMM 794, 470–477.

[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] Hildebrandt, L., Redlich, T., Wulfsberg, J.P., 2020. Persönliche Schutzausrüstung aus der hybriden urbanen Mikrofabrik. ZWF 115 (9), 576–580.

[8] Diez, T., Posada, A., 2013. The fab and the smart city, in: Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction - TEI '13. the 7th International Conference, Barcelona, Spain. 10.02.2013 - 13.02.2013. ACM Press, New York, New York, USA, p. 447.

[9] MacInnis, D.J., 2011. A Framework for Conceptual Contributions in Marketing. Journal of Marketing 75 (4), 136–154.

[10] Torraco, R.J., 2005. Writing Integrative Literature Reviews: Guidelines and Examples. Human Resource Development Review 4 (3), 356–367.

[11] Hees, A., Schutte, C.S.L., Reinhart, G., 2017. A production planning system to continuously integrate the characteristics of reconfigurable manufacturing systems. Prod. Eng. Res. Devel. 11 (4-5), 511–521.

[12] Schuh, G., Scholz, P. Case study on technological applications for production planning and control in the context of industry 4.0. Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020).

[13] Kosturiak, J.Á.N., Gregor, M., 1995. Total production control. Production Planning & Control 6 (6), 490–499.

[14] Garetti, M., Taisch, M., 1999. Neural networks in production planning and control. Production Planning & Control 10 (4), 324–339.

[15] 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.

[16] Oluyisola, O.E., Sgarbossa, F., Strandhagen, J.O., 2020. Smart Production Planning and Control: Concept, Use-Cases and Sustainability Implications. Sustainability 12 (9), 3791.

[17] Albey, E., Bilge, Ü., 2011. A hierarchical approach to FMS planning and control with simulation-based capacity anticipation. International Journal of Production Research 49 (11), 3319–3342.

[18] Nyhuis, P., Münzberg, B., Kennemann, M., 2009. Configuration and regulation of PPC. Prod. Eng. Res. Devel. 3 (3), 287–294.

[19] Wiendahl, H.-P., Breithaupt, J.-W., 1999. Modelling and controlling the dynamics of production systems. Production Planning & Control 10 (4), 389–401.

[20] Sato, R., Tsai, T.L., 2004. Agile production planning and control with advance notification to change schedule. International Journal of Production Research 42 (2), 321–336.

[21] Schuh, G., Prote, J.-P., Luckert, M., Hünnekes, P., Schmidhuber, M., 2019. Effects of the update frequency of production plans on the logistical performance of production planning and control. Procedia CIRP 79, 421–426.

[22] Weng, M.X., Wu, Z., Qi, G., Zheng, L., 2008. Multi-agent-based workload control for make-to-order manufacturing. International Journal of Production Research 46 (8), 2197–2213.

[23] Buckholtz, B., Ragai, I., Wang, L., 2015. Cloud Manufacturing: Current Trends and Future Implementations. Journal of Manufacturing Science and Engineering 137 (4).

[24] Adamson, G., Wang, L., Holm, M., Moore, P., 2015. Cloud manufacturing – a critical review of recent development and future trends. International Journal of Computer Integrated Manufacturing, 1–34.

[25] Liu, Y., Wang, L., Wang, X.V., Xu, X., Zhang, L., 2019. Scheduling in cloud manufacturing: state-of-the-art and research challenges. International Journal of Production Research 57 (15-16), 4854–4879.

[26] Chen, T.-C.T., Wang, Y.-C., 2021. A fuzzy mid-term capacity and production planning model for a manufacturing system with cloud-based capacity. Complex Intell. Syst. 7 (1), 71–85.

[27] Tao, F., Zhang, L., Venkatesh, V.C., Luo, Y., Cheng, Y., 2011. Cloud manufacturing: a computing and service-oriented manufacturing model. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 225 (10), 1969–1976.

[28] Zhou, L., Zhang, L., Sarker, B.R., Laili, Y., Ren, L., 2018. An event-triggered dynamic scheduling method for randomly arriving tasks in cloud manufacturing. International Journal of Computer Integrated Manufacturing 31 (3), 318–333.

[29] Zhang, L., Luo, Y., Tao, F., Li, B.H., Ren, L., Zhang, X., Guo, H., Cheng, Y., Hu, A., Liu, Y., 2014. Cloud manufacturing: a new manufacturing paradigm. Enterprise Information Systems 8 (2), 167–187.

[30] Tao, F., Zhang, L., Liu, Y., Cheng, Y., Wang, L., Xu, X., 2015. Manufacturing Service Management in Cloud Manufacturing: Overview and Future Research Directions. Journal of Manufacturing Science and Engineering 137 (4).

[31] 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.

[32] Ren, L., Zhang, L., Wang, L., Tao, F., Chai, X., 2017. Cloud manufacturing: key characteristics and applications. International Journal of Computer Integrated Manufacturing 30 (6), 501–515.

[33] Wu, D., Greer, M.J., Rosen, D.W., Schaefer, D., 2013. Cloud manufacturing: Strategic vision and state-of-the-art. Journal of Manufacturing Systems 32 (4), 564–579.

[34] Liu, Y., Wang, L., Wang, X.V., Xu, X., Jiang, P., 2019. Cloud manufacturing: key issues and future perspectives. International Journal of Computer Integrated Manufacturing 32 (9), 858–874.

[35] Liu, Y., Xu, X., 2017. Industry 4.0 and Cloud Manufacturing: A Comparative Analysis. Journal of Manufacturing Science and Engineering 139 (3).

[36] Coe, N.M., Dicken, P., Hess, M., 2008. Global production networks: realizing the potential. Journal of Economic Geography 8 (3), 271–295.

[37] Sturgeon, T.J., 2001. How Do We Define Value Chains and Production Networks? *. IDS Bulletin 32 (3), 9–18.

[38] 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), 823–841.

[39] Carvalho, V.M., 2014. From Micro to Macro via Production Networks. Journal of Economic Perspectives 28 (4), 23–48.

[40] Ernst, D., Kim, L., 2002. Global production networks, knowledge diffusion, and local capability formation. Research Policy 31 (8-9), 1417–1429.

[41] Be Isa, J., Zimmermann, T., Scherwitz, P., Reinhart, G. Tactical Production Planning for Customer Individual Products in Changeable Production Networks. Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020).

[42] Davis, A.E., 2018. Global Production Networks and the Private Organization of World Trade. Journal of Economic Issues 52 (2), 358–367.

[43] Kelber, M., Noennig, J.R., Nyhuis, P. Rapid and Long-term Measures for Prevention and Mitigation of Communication Barriers in Production Networks. Proceedings of the 1st Conference on Production Systems and Logistics (CPSL 2020).

[44] Wiendahl, H.-P., Lutz, S., 2002. Production in Networks. CIRP Annals 51 (2), 573–586.

[45] Sturgeon, T.J., 2002. Modular production networks: a new American model of industrial organization. Industrial and Corporate Change 11 (3), 451–496.

[46] Gaalman, G.J.C., Suresh, N.C., 1999. Towards an integration of process planning and production planning and control for flexible manufacturing systems. International Journal of Flexible Manufacturing Systems 11 (1), 5–17.

[47] Gyulai, D., Pfeiffer, A., Monostori, L., 2017. Robust production planning and control for multi-stage systems with flexible final assembly lines. International Journal of Production Research 55 (13), 3657–3673.

[48] Hees, A., Bayerl, C., van Vuuren, B., Schutte, C.S.L., Braunreuther, S., Reinhart, G., 2017. A Production Planning Method to Optimally Exploit the Potential of Reconfigurable Manufacturing Systems. Procedia CIRP 62, 181–186.

[49] Arica, E., Powell, D.J., 2014. A framework for ICT-enabled real-time production planning and control. Adv. Manuf. 2 (2), 158–164.

[50] Rauch, E., Dallasega, P., Matt, D.T., 2018. Complexity reduction in engineer-to-order industry through real-time capable production planning and control. Prod. Eng. Res. Devel. 12 (3-4), 341–352.

[51] 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.

[52] Caridi, M., Cavalieri, S., 2004. Multi-agent systems in production planning and control: an overview. Production Planning & Control 15 (2), 106–118.

[53] Guide, V.D.R., 2000. Production planning and control for remanufacturing: industry practice and research needs. Journal of Operations Management 18 (4), 467–483.

[54] Guide, V.D.R., Jayaraman, V., Srivastava, R., 1999. Production planning and control for remanufacturing: a state-of-the-art survey. Robotics and Computer-Integrated Manufacturing 15 (3), 221–230.

[55] Mcfarlane, D.C., Bussmann, S., 2000. Developments in holonic production planning and control. Production Planning & Control 11 (6), 522–536.

[56] Olhager, J., Wikner, J., 2000. Production planning and control tools. Production Planning & Control 11 (3), 210–222.

[57] Zäpfel, G., Missbauer, H., 1993. Production Planning and Control (PPC) systems including load-oriented order release — Problems and research perspectives. International Journal of Production Economics 30-31, 107–122.

[58] Maccarthy, B.L., Fernandes, F.C.F., 2000. A multi-dimensional classification of production systems for the design and selection of production planning and control systems. Production Planning & Control 11 (5), 481–496.

[59] Starbek, M., Grum, J., 2000. Selection and implementation of a PPC system. Production Planning & Control 11 (8), 765–774.

[60] Alinani, K., Liu, D., Zhou, D., Wang, G., 2020. Service Composition and Optimal Selection in Cloud Manufacturing: State-of-the-Art and Research Challenges. IEEE Access 8, 223988–224005.

[61] Ren, L., Zhang, L., Tao, F., Zhao, C., Chai, X., Zhao, X., 2015. Cloud manufacturing: from concept to practice. Enterprise Information Systems 9 (2), 186–209.

[62] Chen, S., Fang, S., Tang, R., 2020. An ANN-Based Approach for Real-Time Scheduling in Cloud Manufacturing. Applied Sciences 10 (7), 2491.

[63] Lou, P., Hu, J., Zhu, C., Yan, J., Yuan, L., 2021. Cooperation Emergence of Manufacturing Services in Cloud Manufacturing With Agent-Based Modeling and Simulating. IEEE Access 9, 24658–24668.

[64] Guo, L., 2016. A system design method for cloud manufacturing application system. Int J Adv Manuf Technol 84 (1-4), 275–289.

[65] Guo, L., Qiu, J., 2018. Optimization technology in cloud manufacturing. Int J Adv Manuf Technol 97 (1-4), 1181–1193.

[66] Li, W., Zhu, C., Wei, X., Rodrigues, J.J.P.C., Wang, K., 2017. Characteristics analysis and optimization design of entities collaboration for cloud manufacturing. Concurrency Computat Pract Exper 29 (14), e3948.

[67] Simeone, A., Deng, B., Caggiano, A., 2020. Resource efficiency enhancement in sheet metal cutting industrial networks through cloud manufacturing. Int J Adv Manuf Technol 107 (3-4), 1345–1365.

[68] Huang, B., Li, C., Yin, C., Zhao, X., 2013. Cloud manufacturing service platform for small- and medium-sized enterprises. Int J Adv Manuf Technol 65 (9-12), 1261–1272.

[69] Qu, T., Lei, S.P., Wang, Z.Z., Nie, D.X., Chen, X., Huang, G.Q., 2016. IoT-based real-time production logistics synchronization system under smart cloud manufacturing. Int J Adv Manuf Technol 84 (1-4), 147–164.

[70] Bouzary, H., Frank Chen, F., 2018. Service optimal selection and composition in cloud manufacturing: a comprehensive survey. Int J Adv Manuf Technol 97 (1-4), 795–808.

[71] He, W., Xu, L., 2015. A state-of-the-art survey of cloud manufacturing. International Journal of Computer Integrated Manufacturing 28 (3), 239–250.

[72] Moghaddam, M., Nof, S.Y., 2018. Collaborative service-component integration in cloud manufacturing. International Journal of Production Research 56 (1-2), 677–691.

[73] Brydges, T., 2018. “Made in Canada”: Local production networks in the Canadian fashion industry. The Canadian Geographer / Le Géographe canadien 62 (2), 238–249.

[74] DRAYSE, M.H., 2008. Globalization and Regional Change in the U.S. Furniture Industry. Growth and Change 39 (2), 252–282.

[75] Florensa, L.M., Márquez-Ramos, L., Martínez-Zarzoso, I., Recalde, M.L., 2015. Regional versus global production networks: where does Latin America stand? Applied Economics 47 (37), 3938–3956.

[76] Henderson, J., Dicken, P., Hess, M., Coe, N., Yeung, H.W.-c., 2002. Global production networks and the analysis of economic development. Review of International Political Economy 9 (3), 436–464.

[77] Perez-Aleman, P., 2003. Decentralised production organisation and institutional transformation: large and small firm networks in Chile and Nicaragua. Cambridge Journal of Economics 27 (6), 789–805.

[78] Saxenian, A., 1991. The origins and dynamics of production networks in Silicon Valley. Research Policy 20 (5), 423–437.

[79] Jaehne, D.M., Li, M., Riedel, R., Mueller, E., 2009. Configuring and operating global production networks. International Journal of Production Research 47 (8), 2013–2030.

[80] Pomfret, R., 2020. Global Production Networks, New Trade Technologies and the Challenge for International Institutions. Foreign Trade Review 55 (1), 21–41.

[81] Zheng, S., Li, H., Wu, X., 2013. Network resources and the innovation performance. Management Decision 51 (6), 1207–1224.

[82] Hamalainen, M., Karjalainen, J., 2017. Social manufacturing: When the maker movement meets interfirm production networks. Business Horizons 60 (6), 795–805.

[83] Buergin, J., Belkadi, F., Hupays, C., Gupta, R.K., Bitte, F., Lanza, G., Bernard, A., 2018. A modular-based approach for Just-In-Time Specification of customer orders in the aircraft manufacturing industry. CIRP Journal of Manufacturing Science and Technology 21, 61–74.

[84] Dellaert, B.G.C., 2019. The consumer production journey: marketing to consumers as co-producers in the sharing economy. J. of the Acad. Mark. Sci. 47 (2), 238–254.

[85] Hochdörffer, J., Buergin, J., Vlachou, E., Zogopoulos, V., Lanza, G., Mourtzis, D., 2018. Holistic approach for integrating customers in the design, planning, and control of global production networks. CIRP Journal of Manufacturing Science and Technology 23, 98–107.

[86] Coe, N.M., 2012. Geographies of production II. Progress in Human Geography 36 (3), 389–402.

[87] Coe, N.M., Yeung, H.W.-c., 2019. Global production networks: mapping recent conceptual developments. Journal of Economic Geography 19 (4), 775–801.

[88] KRÄTKE, S., 2014. How manufacturing industries connect cities across the world: extending research on ‘multiple globalizations’. Global Networks 14 (2), 121–147.

[89] MacKinnon, D., 2012. Beyond strategic coupling: reassessing the firm-region nexus in global production networks. Journal of Economic Geography 12 (1), 227–245.

[90] Mahutga, M.C., 2014. Production Networks and the Organization of the Global Manufacturing Economy. Sociological Perspectives 57 (2), 229–255.

[91] Mella, P., 2019. The ghost in the production machine: the laws of production networks. K 48 (6), 1301–1329.

[92] Yeung, H.W.-c., 2015. Regional development in the global economy: A dynamic perspective of strategic coupling in global production networks. Regional Science Policy & Practice 7 (1), 1–23.

[93] Yeung, H.W.-c., 2020. Regional worlds: from related variety in regional diversification to strategic coupling in global production networks. Regional Studies, 1–22.

[94] Kuehnle, H., 2007. A system of models contribution to production network (PN) theory. J Intell Manuf 18 (5), 543–551.

[95] Klooster, D., Mercado-Celis, A., 2016. Sustainable Production Networks: Capturing Value for Labour and Nature in a Furniture Production Network in Oaxaca, Mexico. Regional Studies 50 (11), 1889–1902.

[96] Pannok, M., Finkbeiner, M., Fasel, H., Riese, J., Lier, S., 2020. Transformable Decentral Production for Local Economies with Minimized Carbon Footprint. ChemBioEng Reviews 7 (6), 216–228.

[97] Prause, G., Atari, S., 2017. On sustainable production networks for Industry 4.0. JESI 4 (4), 421–431.

[98] Sandkuhl, K., Smirnov, A., 2018. Context-oriented Knowledge Management in Production Networks. Applied Computer Systems 23 (2), 81–89.

[99] Branding, J.-H., Basmer-Birkenfeld, S.-V., Redlich, T., 2019. Using Open Production Sites for Supporting New Ways of Corporate Innovation, in: Redlich, T., Moritz, M., Wulfsberg, J.P. (Eds.), Co-Creation. Springer International Publishing, Cham, pp. 21–35.

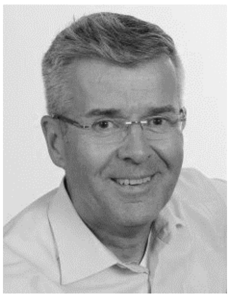
[100] Redlich, T., Moritz, M., 2016. Bottom-up Economics. Foundations of a Theory of Distributed and Open Value Creation, in: Ferdinand, J.-P., Petschow, U., Dickel, S. (Eds.), The Decentralized and Networked Future of Value Creation. Springer International Publishing, Cham, pp. 27–57.

[101] Redlich, T., Wulf, S., Moritz, M., Buxbaum-Conradi, S., Krenz, P., Wulfsberg, J., 2015 - 2015. The Strategy of Openness in industrial production, in: 2015 Portland International Conference on Management of Engineering and Technology (PICMET). 2015 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA. 02.08.2015 - 06.08.2015. IEEE, pp. 302–309.

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.

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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Potential Analysis of Flexible Small Series Production of Spare Parts by Direct Polymer Additive Tooling

Günther Schuh[1], Georg Bergweiler[1], Falko Fiedler[1]
Daniel Neumann[1], Lennart Heetfeld[2], Philipp Bickendorf[1*]
*Corresponding author (philipp.bickendorf@rwth-aachen.de)
[1]Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Aachen, Germany
[2]Chair of Production Engineering of E-Mobility Components (PEM), RWTH Aachen University, Aachen, Germany

Abstract

In recent decades, there has been an increasing creation of vehicle model variants and product individualization for customer needs. There is a need for a greater variety of spare parts which is leading to high requirements on their logistics. As a result of this trend, automotive production companies have to ensure the provision of many different spare parts as part of their post-fulfillment obligations. However, the surplus spare parts stock will increase the operation cost of the company. This implies the need for a more flexible approach to spare parts production and provision to improve profitability compared to existing process chains. In this research paper, the potential of flexible spare parts production and provision is discussed. For this purpose, the state of spare part production technologies and existing approaches for spare part service are analyzed regarding their technological characteristics, flexibility as well as cost structure. Following this, an approach for flexible production for spare parts with regard to its potential is analyzed and demonstrated based on three exemplary automotive use cases using additively manufactured production resources. In the casting use case, a Fused Filament Fabricated (FFF) additively manufactured speedometer screw model will be used as a sand casting form. The production of deep-drawn car body parts using polymer-based FFF forming tools is investigated in the second use case. Lastly, the production of ignition distributor caps by PolyJet Modelling (PJM) manufactured Injection Molding (PUR-RIM) molds is presented.

Keywords: Additive Tooling; Rapid Tooling; Spare Parts; Classic Cars; Small Series Production; Production System; Production Concept, Production Equipment; Automotive Production

1. Introduction

Enabling technologies have diverse technology-inherent annual quantity spectra and allow economical production of parts of different complexity in small series. In recent years, spare parts production and provision have gained importance and are becoming increasingly relevant in all production industries. For example, a typical problem in the automotive industry is the build-up of excessive spare parts inventories. This is done based on of flat-rate surcharges on forecast requirements, but not on actual demands. For this reason, Original Equipment Manufacturers (OEM) and suppliers tie up capital in an unprofitable way due to inflexible, conventional production of spare parts [1]. Similar problems can be seen in the case of classic cars, too. Despite an increasing number of registrations of classic cars, the demand for spare parts is severely limited. In addition, in many cases the original tools for manufacturing parts no longer exist, so that spare parts production and supply under conventional aspects prove to be unprofitable [2]. To be able to offer high quality spare parts at competitive prices, it is therefore necessary to identify the potentials of enabling production technologies, such as Additive Manufacturing (AM), that can be used to rebuild required tools or the part itself. For this purpose, production technologies must be analyzed and cost drivers identified. To enable a price evaluation of parts to be produced, a classification of parts must be created, which can be linked to the desired market prices. The identified cost drivers and the approximation of market prices can be used to evaluate the economic efficiency of the flexible production of automotive spare parts. Despite high expected cost of AM and uncertainties of its technological development, AM is promising a reduction of lead times, tool and storage cost as well as an increase in production flexibility [3].

Therefore, the underlying research question arises how significant the potential is for the use of Direct Polymer Additive Tooling (DPAT) in the production equipment manufacturing for spare parts and how potentials can be used. Hence, the research hypothesis of this paper is that Rapid Additive Tooling (AT) has a high potential for a cost- and time-effective production and provision of automotive spare parts with automotive-typical quality. Furthermore, a systematical scheme for the application of AT for production equipment in small series production can be conceptualized. To investigate this hypothesis, the current state of research is discussed. Deficits of existing approaches with regard to the industrial requirements of spare parts are described and potential solutions are presented. Based on this, a potential analysis scheme for the use of AM in production equipment manufacturing is conceptualized as well as demonstrated by three automotive use cases. [4,5]

2. State of the Art

Production of Spare Parts: A production system describes the appropriate combination of resources (material, energetic and spatial, as well as personnel and financial), information and competences for the manufacturing of products of a certain type and quantity at acceptable costs with a specified quality. Production can be understood as any combination of production factors, which thus encompasses the operational performance process and its corporate potential for production tasks can be derived [6]. This understanding of production systems has also been established in the automotive industry [7]. Due to the aftersales services, where typically about 22 years including the 10-year legal spare parts supply obligation elapse from the Start of Production (SOP) of the vehicle models to the End of Delivery Obligation (EDO), a production system is essential. In this system, spare parts are produced during the production phase of the models to compensate for early defects, followed by a consolidation phase with fluctuating demand for spare parts and a degeneration phase. Production is usually carried out by spare parts suppliers on the serial equipment used until then, i.e. with the help of equipment that is manufactured by casting and chipping, which are planning-time- and cost-intensive. After End of Production (EOP), the stocking of high-quality spare parts for the vehicle lifetime has to be foreseen, with the main difficulty being the long-term planning horizon with high uncertainties. [8]

Spare parts production follows the typical production understanding of the automotive industry with decentralized material provision, production and logistics, whereby lot sizes are often planned according to general ordering algorithms of business administration [9,10]. To counteract the uncertainties in planning, there is a move towards lean spare parts management as a logistical trade-off between provisioning, capital commitment and costs [11]. Spare parts logistics as the second aspect of spare parts provision alongside production is responsible for the supply of materials and information, procurement from the supplier as well as warehousing and finally provision at the point of need. In this context, replacement and consumable parts are scheduled and procured according to the same principle, without considering changes in defect or consumption characteristics. This often results in oversized safety stocks. [1] Expert statements underline the requirements for the quality of spare parts in the sense of the automotive understanding of production. Despite low unit demand and the desire for competitive costs, OEMs still rely on extensive technology exchange for R&D and internal storage of high volumes of spare parts. [12] This traditionally grown spare parts supply results in a high inventory capital and high investment costs in conventional tools (e.g. by casting and chipping) with an inflexible provision and strongly physical understanding, whereas current trends towards new value creation approaches such as the copyright on product and production data are hardly considered. The result is a low service level, even though the spare parts business holds enormous potential for success for OEMs. [10,8] The lack of flexibility in the provision of spare parts is a problem for classic cars, too. In contrast to the situation with current models, there are often no longer any series tools for classic cars and, due to the lack of digital product and production data, these cannot be easily reproduced. [8] Furthermore, the scattering of vehicles over the course of time makes the logistics processes of spare parts supply for classic cars even more difficult. Hence, the resulting costs from a conventional point of view for spare parts for classic cars are even higher than they are in the context of general after-sales services.

Enabling Production Technologies: In addition to conventional machining, enabling production technologies with great relevance have emerged as a result of technological developments. In the following, AM will be examined in greater detail. AM describes generative 3D printing in which objects are produced directly on the basis of Computer Aided Design (CAD) models. Manufacturing is accomplished by the layer-by-layer automated application of liquids, powder and filament material to produce true-to-scale 3D physical objects without the use of part-specific tools [13,14]. The printing process itself is thereby automatically determined by the computer based on the CAD, enabling the implementation of an end-to-end data process chain [14]. Meanwhile, all printers use the basic layer buildup approach, but the relevant parameters of the process for differentiation represent the materials used [13]. The build-up principle of AM thus allows complex product designs with material-dependent product properties and high dimensional stability, whereby variable part costs are mainly dependent, when neglecting the amortization of investment costs, on post-processing and material costs [15–17]. The most widespread AM technology is FFF, a process in which thermoplastic is heated above its liquefaction point and applied through a nozzle in layers to the object geometry. Advantages of FFF are that the use of support materials makes it possible to create very complex geometries that cannot be produced conventionally, e.g. by injection molding, and a very wide industrial and private distribution. Disadvantages are a lower surface quality and poorer mechanical properties of parts. PJM, a further widely used process, relies on photopolymers and thus achieves a higher degree of accuracy and greater variability of printing properties, which, however, also results in higher costs. Accordingly, the advantages of AM are short lead times after part design as well as rapid production of parts in the printing process and the elimination of some time-consuming process steps. In the following, DPAT as a subcategory of AM is of particular importance in the context of production tooling for the spare parts production and will be examined in more detail below on the basis of initial approaches from research and industry. [14]

Existing Research Approaches: The AM technology presented can be used in a variety of ways for spare parts provision in automotive industry due to the large number of different process and parameter combinations. In the following, an overview of some research approaches and industrial use cases of AM and AT for small series production and spare parts provision will be presented. The technical possibility and economic feasibility of AM and AT applications depend on production volume, part size, material cost and complexity. DPAT can produce equipment quickly and cost-effectively. This methodology has already been widely used in industry for several years and can be applied for the small-batch mold production of interior parts [18]. AT is cost effective with plastic injection molding of volumes around 1,000 parts [19]. Shaping tool elements can also be substituted by AT and research has shown potential its regarding the use in deep drawing, stamping and bending as well [20]. Additively manufactured tool inserts can be integrated into a modular pillar set out of conventional tooling steel [21]. In this way, deep drawing inserts can be manufactured for the use in small series of sheet metal parts with internal resources [5]. Furthermore, intensive research is being conducted into the further development of AM materials. E.g., material combinations of FFF filaments enable the production of parts with a stiffness comparable to that of aluminum. [22] The logic of substituting shaping tool elements can be applied not only to other production equipment (such as welding jigs or assembly aids), but also to other production technologies (such as AM of sand cast model and core box model). Research for large-scale industrial applications is being carried out, e.g. at Volkswagen's Portuguese plant, where most assembly jigs are being produced internally with the help of AM instead of external partners. Further approaches for e.g. fixture design specifically for electro mobility are also being researched and continuously developed [23–25]. In addition to the existing approaches for AT, reference should also be made to the direct printing of spare parts. OEMs are producing e.g. mechanical parts in small batches, particularly in the luxury segment: Bugatti produces AM brake calipers from titanium for series equipment. For spare parts production, the Volkswagen Group started to produce individual AM of release levers for the clutch as spare parts e.g. for the Porsche 959. This results in a broad field of initial application areas for AM/AT in the automotive industry for series and spare parts requirements [26].

3. Analysis of the Spare Parts Supply Chain

Deficits of Supply Chain: In the following, the currently prevailing disadvantages of spare parts supply will be presented and the areas of application of possible new technologies will be shown. The challenge focused on in this paper is the integration of AT into prevailing structures. DIN 24420 is used to enable a clear distinction of parts, assembly groups as well as complete products as spare parts [27]. They are intended to maintain the function of a vehicle and can include both worn and missing parts [28]. Regarding demand patterns during the production, compensation and degeneration phases, a distinction of spare parts is made between wear and defective parts [28]. Wear parts are generally in high demand, which means that even after the EDO, procurement is usually possible without any problems. Wear parts are simple parts such as rubber bushings or brake pads [29]. The predictability and availability of wear and non-wear parts is good since they are mass-produced, statistics from previous models can be used as a reference. Thus, the post-fabrication options are usually unproblematic. In contrast, defective parts have a low request and the defect time is not well predictable ('spikes'). Defective parts are mostly parts which are highly innovative and complex. Electrical parts are mostly affected, due to the increasing amount of electric parts in the vehicle [29,30]. The production period is also usually very short, as these parts are expensive and therefore rarely kept in stock [31]. For this reason, attempts are made to replace defective parts with carry on parts or substitution parts [32]. From today's perspective, it can be seen that the older the vehicle or the higher the mileage, the more likely it is that not only mechanical parts but also body or interior parts will fail. The current spare parts situation varies depending on age, number of vehicles built and OEM. The supply of wear parts is fine. However, it can be seen that after EOP and EDO, there is a time when real demand for classic cars is met by other means. First of all, remaining stocks are used up or 'wrecked cars' are exploited. Conversions to non-original parts are also carried out to keep cars roadworthy. Only after these variants have been used up, a higher demand from the OEM becomes apparent. As an alternative, spare parts are also produced by private individuals or even interest groups (more common) to counteract bottlenecks.

Besides these general spare part supply activities, the classic spare part production starts normally after EDO respectively even further from the EDO. Vehicles must be at least 30 years old to be classified as a classic car and depending on era as well as category (mass/premium vehicle), the complexity of the vehicles and parts differ [33]. However, electrification was very limited in classic cars production and defective parts therefore are mainly mechanical or body parts. [33]. Currently, subcontractors are founded that take care of the spare parts procurement of classic cars. Here, following the above-mentioned general supply approach, the OEM buys up remaining stocks from suppliers or produces new parts with the help of existing tools. However, as this is very expensive and customers demand a high quality, more complex parts are slow to be reproduced. At some point, these supply solutions are no longer technologically and economically worthwhile and are provided for brand management and customer/fan loyalty [12]. In Germany, classic cars spare part demand can be forecasted well as the number of active cars is completely known [34]. For classic cars, a distinction must be made between visible and functional parts. Critical functional parts may deviate from the original part if there is no other way to keep the car roadworthy. Visible parts (non-critical) are usually not accepted if the quality deviates from original parts. Some wear parts for classic cars, such as shaft bearings or brake pads, are even still available, as they are widely used in other industries as well. Substitution is therefore also a possibility to fulfil one's need for spare parts.

Analysis of Enabling Production Technologies: Various approaches regarding e.g. process optimization, sustainability and the simulation of supply and demand of spare parts with the help of AM has been object of research so far, but no general approach for a production system for AT has been developed [4,3,20]. As mentioned in chapter 2, a distinction must be made between AM and AT. In AT, a shaping tool which is additively manufactured is used to form or cast the spare part. With direct AM, there are completely different materials or materials with comparable mechanical characteristics available to produce the spare parts. Choosing the use case specific combination of technology, process and material is important. Following this, potential advantages are reduced part or tool costs, shorter lead times, a high potential regarding automation as well as maximum potential regarding adding geometric complexities. AM is highly flexible and reworking parts is usually less compared to conventional production technologies.

The challenges are to increase process stability and implement the currently prevailing build volume limitations of AM in the best possible way. For classic car spare parts, the quality of visible and functional parts must also correspond to that of the original part. A closer look at the product life cycle reveals that AM/AT have applications in every phase. Due to technological boundaries such as print time and layer height, the technological and economic potentials are so far mainly enabled in small annual quantities. These mostly include the prototype, product development and spare parts sectors. Nevertheless, there is no systematic classification in automotive production concepts so far. [10]

4. Flexible Spare Parts Provision

Production Characteristics: In the following, use case specific characteristic values will be described and used as a comparison to conventional tooling. For the successful establishment of a flexible production of spare parts, it is essential to consider the potential of a manufacturing task within the overall corporate value creation context. A five-stage procedure can be used for this purpose: 1) Analysis of the spare parts portfolio and the need for production equipment 2) Challenging production processes and their potential for DPAT 3) Analyzing business cases 4) Derivation of AM strategy 5) Piloting of approach, validation and rollout [10]. With the help of six production characteristics defined in Table 1, a production company can assess such procedures.

Table 1: Production characteristics

Characteristic	Short description
Process	Process considers all elements of the service creation process that are necessary for part production (development, design and production of part & tool).
Materials	Materials include characteristics resulting from part and tool material choices. If the original material is not available, an at least equivalent quality level is to be aimed for.
Function	Function records which requirements exist for the part, e.g. functional requirements in terms of mechanical, thermal and chemical resistance or visual and haptic impressions.
Geometry	Geometry assesses part shape and dimensions. The focus is on geometric elements that increase the design effort or impose restrictions in terms of production technology.
Costs	The costs are divided into personnel costs, operating material costs, material costs and capital costs for the part as well as any necessary tools. It is assumed that the provision time of the tools resulting from the production process correlates with the costs.
Tool life	The expected quantity to be produced is considered within this characteristic. A tool is worn when the required part quality or dimensional accuracy can no longer be achieved.

On top of these six exemplary characteristics, further aspects as for example the original material, small batch quantities, reduction of tooling costs, use of conventional manufacturing process, and existing production facilities might need to be considered for a corporate-individual potential analysis as well. The exemplary outputs of this procedure will be described in the following and allows an evaluation of AT for flexible small series spare parts production based on these experiences. The main aspects of the potential of AT are summarised below.

Use Case 1 - Sand Casting by AM Mold Models: Typical cast components include engine components, transmission components, brake discs and housing components. The part manufactured in this use case is a speedometer screw (production date 1933) which is located near the wheelhouse and is mounted on a drive shaft that transmits the signal for display in the speedometer scale (cf. Figure 1). This shaft cover (defective part) encloses a drive shaft in which a magnetic disk rotates to indicate the vehicle's speed. At first, the original part is analyzed and the dimensions are determined by optical 3D scanning. When the manufacturing process was initially carried out, deriving the dimensions was difficult due to the deformation of the original component and required increased rework in reengineering a 3D model during CAD design. This is a common challenge for the reengineering of classic car spare parts. The next step is to prepare the FFF of the mold models. In sand casting, these mold models will be pressed into oil sand to leave the part mold for casting in it. The mold models are manufactured out of a wax filament with printing parameters that allow low material consumption, sufficient stiffness as well as an accurate and smooth surface. Using the AM mold models, the sand mold is prepared for investment casting. After removing the mold models, the aluminum will be poured in the sand mold by means of runners. Finally, the cast parts are subsequently cleaned and reworked. In the process, the sprue is removed as well as holes and a thread are inserted, which cannot be produced by sand casting. Because the molding model can be removed prior to the sand casting, it can be used as often as desired to produce a sand mold. The surface quality can be adjusted according to the AM quality and the degree of post-processing of the mold models and is therefore accordingly good. The manufacturing costs include the material, labor and machine costs associated with the part as well as the mold form. The detailed cost of eight parts is shown in Table 2. In terms of potential evaluation, it can be summarized that the tooling costs are significantly lower than for milled permanent metal models, and undercuts and internal geometry that normally generate significant additional work (core and core box) do not pose any problems when designed for AM and casting. Therefore, DPAT for sandcasting shows potential for small series production concerning the production characteristics process, geometry, cost and tool life from chapter 4.1.

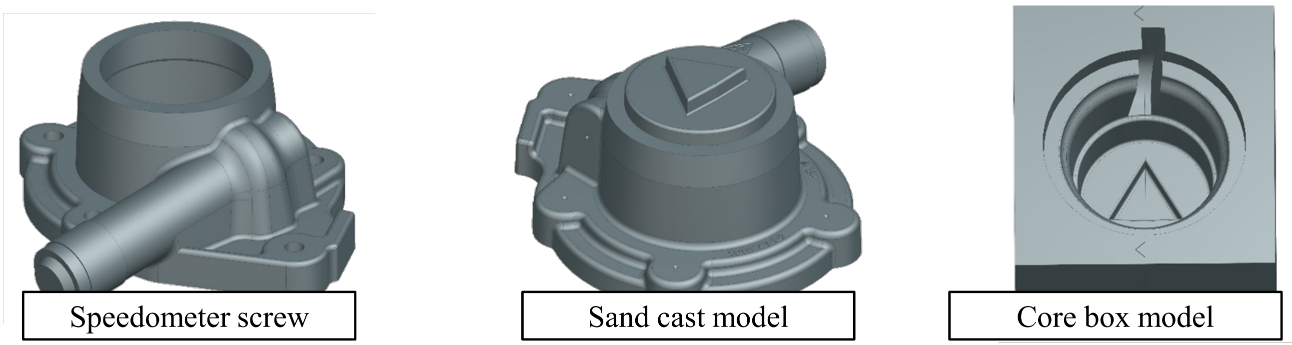


Figure 1: Speedometer screw and AM sand casting models

Table 2: Exemplary cost breakdown

Cost breakdown sand casting	Time needed [h]	Hourly rate [€/h]	Series costs [€]
3D scan (Labor + hour rate)	2.00	70.00	140.00
CAD design	20.00	50.00	1,000.00

1. Filament (PLA & Support) 2. Personnel (Plant operator) 3. Machine hour rate (Ultimaker S5)	 1.00 2.50	 25.00 18.40	6.00 25.00 46.00
1. Molding sand (Oil-based) 2. Aluminum (AlMgSi) 3. Labor (Plant operator)	 12.00	 25.00	80.00 1.79 300.00
Total cost for a small series of eight parts			1,598.79

Use Case 2 - DPAT for Deep Drawing: The part to be manufactured is a seatbelt retainer plate (defective part), which is installed on the inner B-pillar to reinforce the seatbelt attachment area. This is a typical crash-relevant small car body shell part out of sheet metal with small tolerances on joining and functional elements (+/- 0.5 mm) (cf. Figure 2) [5]. Following the design of the sheet metal part, the CAD model is extended by a seam for the blank holder [35]. A formability simulation with the help of Ansys determines whether the component is at risk of cracking, wrinkling or compression during the forming process. The layout and design of the tool can be derived with the help of Computer aided sheet metal forming add-ons. A FEM simulation was also carried out for the mold, which enables a CAD design optimized for the application. An influence that has not yet been fully quantified is the elastic-plastic deformation of the tool material. Based on experience from previous experiments, Polylactic Acid (PLA) is chosen for the FFF of the tool because of its good cost-per-stiffness value as well as easy processability. Due to the high quality of the print results achieved with PLA, post-processing was limited to the removal of the build plate adhesion as well as a reaming of the holes for the dowel pins. For the integration into the forming press, the tools were mounted on a pillar set. For a small series of 27 components, 16 blanks of 1 mm DC04 deep drawing steel and 11 blanks of 1 mm 5754 aluminum were laser cut. After deep drawing with no lubrication, the parts need to be cut by 3D laser cutting to achieve the final shape. The part has various geometry elements such as stiffener beads or joining flanges which are common for sheet metal parts. The part's outer sides have 10 mm radii, whereas the maximum drawing depth is 18.5 mm. The tool set consists of a punch, die and blank holder, which are printed one after the other due to build space constraints of the Desktop printer used. With the help of AT, a significant reduction in costs similar to the first use case can be achieved. This can be justified with a shorter tool life due to the material as well as friction and wear. The tool therefore shows potential regarding small series deep drawing applications by considering the production characteristics material, function, geometry, costs and tool life.

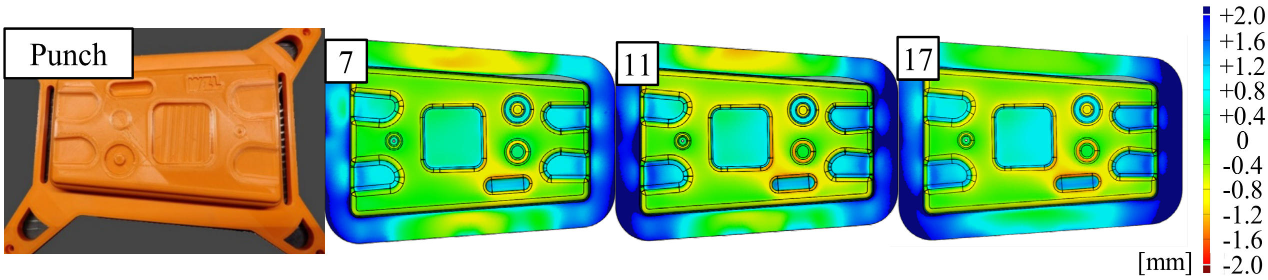


Figure 2: AM deep drawing tool and dimensional accuracy of sheet metal parts after forming operation 7, 11, 17

Use Case 3 - DPAT for Injection Molding: The part to be manufactured is an ignition distributor cap (wear part) produced by AM PUR RIM molds (cf. Figure 3). Scanning of the original component and reengineering is performed similarly to the first use-case. In the next step, a CAD design of the mold tool is created. The final injection molding then is printed using PJM. For the creation of a smooth surface, the cavity needs to be grinded in the post-processing. With the mold tool printed, the next step is to cast the ignition distributor cap using PUR RIM and subject it to post-processing, such as removal of the sprue and venting, before installation. During PUR RIM, a cast resign is used for the application of thin-walled and impact-resistant components. The material requirements result from the use in the engine compartment under corresponding thermal, mechanical and weathering influences. In addition, there is the requirement of a high optical as well as haptic quality with fine geometries for the component, which must be met in PUR RIM. Typical production of parts like this in the automotive industry takes place in annual quantities of over 10,000 and with short cycle times of about one second. In contrast, PUR RIM with the AM mold enables cost-efficient small-batch casting of the components with a high degree of design freedom, but at the same time also a limited range of application with regard to thermal and mechanical loads on the printed molds in PUR RIM. To meet these requirements for the mold, the two halves of the pattern used to manufacture the distributor cap were produced using PJM. By using release agents in PUR RIM, the mold can ultimately be reused for up to 50 cast parts before the required quality can no longer be maintained by the printed mold. For further evaluation, critical aspects are temperature resistance and surface quality of the mold configuration. Regarding the six production characteristics from chapter 4.1, the small series potential can be derived. Thus, the use of AM in injection molding is already established in some industrial applications.

Potential of AT: According to the preceding use cases, AT offers some advantages despite a lack of experience, which are summarized below. AT can usually meet the high requirements for the originality of components better than AM, especially regarding the use of material. Depending on the AM process, only a certain selection of materials can be processed. AT reduces the economic risk compared to milled steel tools, as tooling costs and provision time are minimized. It is more profitable than AM in small to medium quantities, especially quantities of 8-15 parts can be represented well using AT. The disadvantage of AM is that part costs remain almost constant across the quantities. Experience with AT can be transferred to many other use cases such as prototyping or product development. Existing production equipment and facilities can be used. Furthermore, the conventional production process is used, in which empirical knowledge and human resources are already available. This facilitates possible homologation or certification procedures.

5. Conclusion

Summary: In this paper, the potential of DPAT in the spare part production and supply is analyzed. This is done by analyzing existing spare part production chains as well as presenting three use cases with additively manufactured tools. The examined use cases are evaluated on the basis of six production characters regarding Process, Materials, Function, Geometry, Costs and Tool life for both, wear parts and defective parts. In the first use case, a surface optimized FFF-ABS sand cast model as well as core box model will be used as

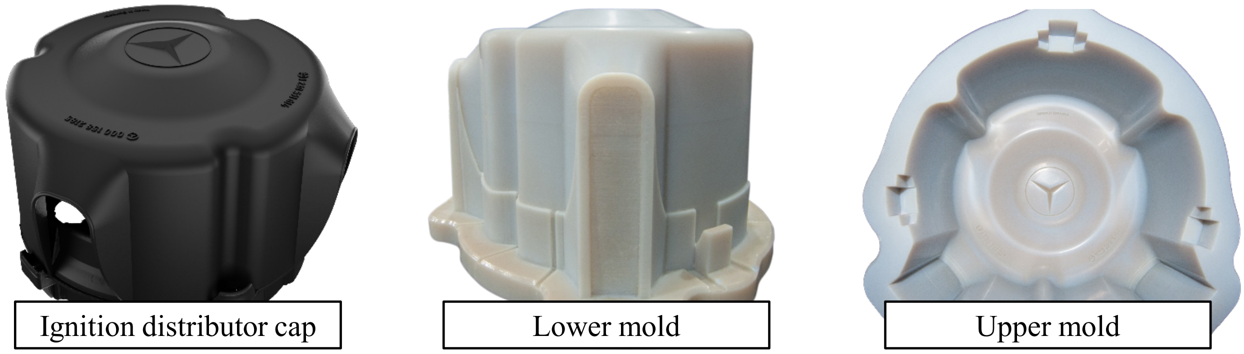


Figure 3: Ignition distributor cap (spare part) and AM injection molds

production equipment for the small series sand casting of a speedometer screw. The production of cast molds is making use of the high degree of geometrical freedom of AM. In the second use case, a small series of small sized car body parts are deep drawn with an AT setup out of FFF PLA. The tool life is sufficient for the production of 17 seatbelt retainer plates out of 1 mm DC04 steel. Stiffness and wear resistance of such tools is limited so that deformation and wear play a major role. FFF is suitable for the production of forming tools for small series and spare parts applications, because of the low investment, process and material costs. A distributor ignition cap is manufactured with injection molding (PUR RIM) in the third use case. The PJM of Digital ABS material was used to manufacture the upper and lower injection molds needed for the production. Due to the complex geometries and the time-consuming reverse engineering, this is the most complex use case in terms of time, costs and requirements. Overall, AT shows promising results in specific use cases for the improvement of spare parts production and supply. Especially for the supply of classic cars, which are not supplied by the OEMs after EDO, AT can speed up production and reduce costs of small series. So far, AM has shown potential regarding the direct AM of spare parts mainly due to the non-necessity of shaping elements and tools. But in the case of AT, components can be cast or formed in the original materials in conventional quality. In addition, the mechanical characteristics of additively manufactured production equipment can be adapted according to the specific local load with the help of a broad variety of AM materials. Based on the demonstration of three use cases, a high potential for the use of DPAT for the production of spare parts can be observed. The systematic approach with an organized framework regarding the production characteristics might be a promising approach for spare parts provision.

Limitation: The scope presented is limited to the analysis of conventional production technologies with the great importance in spare parts supply such as metal casting, deep drawing and injection molding. By replacing conventional manufacturing of production equipment with A, process limits of the tooling configurations are changed. Thus, new competencies have to be built up with respect to the shifted process boundaries. So far, AM is only used for the production of mechanical parts such as housings, caps or car body parts. A potential analysis of the production of electrical, electronical and mechatronic parts needs to be done.

Outlook: Further production technologies in which AM show potential are hydroforming, rubber pad forming, investment casting and bending. In hydroforming, parts of the highly complex tooling set can be manufactured by AT. Elastic AM materials are used in rubber pad forming to print the rubber pad, which can be used flexibly for different forming geometries compared to deep drawing. In investment casting, filaments can be used to print the lost form models, which evaporate in contact with the casting material in a similar way to the wax previously used for this purpose. The design flexibility of AM allows complex geometries to be created. In other forming processes such as bending, various high-stiffness AM materials can be used to replace series steel tools. Furthermore, the tool deformation prediction is further investigated. For this, a database by evaluation data from further experiments can be created. The results are transferred to new geometries on different parts to accelerate the design. The portfolio can be expanded regarding other parts such as profiles. AM is also enabling potential regarding the connection with digital business models. Spare parts can be produced on demand with the help of AM/AT process chains. For the confirmation of the industrial suitability, various use cases can be integrated in real production environments.

Acknowledgements: We wish to acknowledge the BMWi for funding the project ‘LeSS’.

References

[1] Pawellek, G., 2016. Integrierte Instandhaltung und Ersatzteillogistik. Springer, Berlin, Heidelberg.

[2] Hamill, S., 2020. 3D Printed Car Parts: The Savior of Classic Cars | 3DSourced.

[3] Knofius, N., van der Heijden, M.C., Zijm, W.H.M., 2019. Moving to additive manufacturing for spare parts supply. Computers in Industry 113, 103134.

[4] Bergweiler, G., Heinze, F., Lichtenthäler, K., Bickendorf, P., 2020. Ressourceneinsparung durch additiv gefertigte Umformwerkzeuge und Schweißvorrichtungen. ATZ Produktion 7 (1).

[5] Schuh, G., Bergweiler, G., Fiedler, F., Bickendorf, P., Schumacher, P. Small Series Production and Geometric Analysis of Sheet Metal Car Body Parts Using Forming Tools Made of FFF PLA. IEEM, 156–160.

[6] Westkämper, E., 2006. Einführung in die Organisation der Produktion. Springer, Berlin, Heidelberg.

[7] Kampker, A., Bergweiler, G., Hollah, A., Bickendorf, P., Hoffmann, F., 2019. Development of a framework for the flexibility assessment of automotive production concepts. Procedia CIRP 81, 34–39.

[8] Klug, F., 2018. Logistikmanagement in der Automobilindustrie. Springer, Berlin, Heidelberg, 528 pp.

[9] Biedermann, H., 2008. Ersatzteilmanagement. Springer, Berlin, Heidelberg, 159 pp.

[10] Geissbauer, R., Wunderlin, J., Lehr, J., 2017. The Future of Spare Parts is 3D, 2017.

[11] Guo, D., Yu, X., 2018. The Lean Management of Spare Parts in Automotive Manufacturing. MATEC 214, 4005.

[12] Bickendorf, P., 2020. Innovationspotenziale für die Ersatzteilproduktion bei deutschen OEM.

[13] Dudek, P., Rapacz-Kmita, A., 2016. Rapid Prototyping: Technologies, Materials and Advances. Archives of Metallurgy and Materials 61 (2), 891–896.

[14] Fastermann, P., 2016. 3D-Drucken. Springer, Berlin, Heidelberg.

[15] Atzeni, E., Iuliano, L., Minetola, P., Salmi, A., 2010. Redesign and cost estimation of rapid manufactured plastic parts. Rapid Prototyping Journal 16 (5), 308–317.

[16] Franchetti, M., Kress, C., 2017. An economic analysis comparing the cost feasibility of replacing injection molding processes with emerging additive manufacturing techniques. Int J Adv Manuf Technol 88 (9-12), 2573–2579.

[17] Schröder, M., Falk, B., Schmitt, R., 2015. Evaluation of Cost Structures of Additive Manufacturing Processes Using a New Business Model. Procedia CIRP 30, 311–316.

[18] Schuh, G., Bergweiler, G., Lukas, G., Abrams, J.A., 2020. Feasibility and Process capability of polymer additive injection molds with slide technology. Procedia CIRP 93, 102–107.

[19] Sedacca, B., 2011. Hand built by lasers: Additive layer manufacturing. Engineering & Technology 6 (1), 58–60.

[20] Nakamura, N., Mori, K.-i., Abe, Y., 2020. Applicability of plastic tools additively manufactured by fused deposition modelling for sheet metal forming. Int J Adv Manuf Technol 108 (4), 975–985.

[21] Schuh, G., Bergweiler, G., Bickendorf, P., Fiedler, F., Colag, C., 2020. Sheet Metal Forming Using Additively Manufactured Polymer Tools. Procedia CIRP 93, 20–25.

[22] Bergweiler, G., Fiedler, F., Kampker, A., Lichtenthäler, K., 2019. Additiv gefertigte Umformwerkzeuge im Prototypenbau. Umformtechnik 2019 (3), 18–19.

[23] Kampker, A., Bergweiler, G., Hollah, A., Lichtenthäler, K., Leimbrink, S., 2019. Design and testing of the different interfaces in a 3D printed welding jig. Procedia CIRP 81, 45–50.

[24] Lichtenthäler, K., Bergweiler, G., Kampker, A., Hollah, A., Leimbrink, S. AM for Cost Efficient Hybrid Welding Jigs, 559–571.

[25] Schuh, G., Bergweiler, G., Lichtenthäler, K., Fiedler, F., La Puente Rebollo, S.d., 2020. Topology Optimisation and Metal Based AM of Welding Jig Elements. Procedia CIRP 93, 62–67.

[26] Volkswagen AG, 2018. Bremssattel und Räder? Jetzt aus dem 3D-Drucker!

[27] 1976. DIN 24420-1:1976-09: Ersatzteillisten; Allgemeines.

[28] Dombrowski, U., Malorny, C., Kolshorn, R., Blöcker, T., 2008. Teiledienst, 35–124.

[29] Nyhuis, P., 2008. Beiträge zu einer Theorie der Logistik.

[30] Roland Berger Strategy Consultants, 2010. Industrieservices in Deutschland.

[31] Livingston, H.C., 2002. GEB1: Diminishing Manufacturing Sources and Material Shortages (DMSMS) Management Practices.

[32] Dombrowski, U., Fochler, S., Malorny, C., 2020. After Sales Service. Springer, Berlin, Heidelberg.

[33] 2020. FIVA Technical Code 2020: Protecting, Preserving and Promoting the use of Historic Vehicles.

[34] Verband der Automobilindustrie e.V., 2021. Der Deutsche Oldtimer Index.

[35] Dietrich, J., 2018. Praxis der Umformtechnik. Springer, Wiesbaden.

Biography

Prof. Dr.-Ing. Günther Schuh (*1958) is head of the Chair of Production Engineering at WZL at RWTH.

Dr.-Ing. Georg Bergweiler (*1981) is head of the department Vehicle Production at WZL at RWTH.

Falko Fiedler, M.Sc. (*1988) is group lead Body Shop at WZL at RWTH.

Daniel Neumann, B.Sc. (*1996) is a student research assistant at WZL at RWTH.

Lennart Heetfeld, M.Sc. (*1991) is a research associate at PEM at RWTH.

Philipp Bickendorf, M.Sc. (*1992) is a research associate at WZL at RWTH.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Feasibility Analysis of Entity Recognition as a Means to Create an Autonomous Technology Radar

Günther Schuh[1], Jan Hicking[1], Max-Ferdinand Stroh[1], Justus Benning[1], Clinton Gnanaraj Charles[1]

[1]*FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany*

Abstract

Keeping up to date with the latest technology trends is crucial task for manufacturing companies to remain successful on a globally competitive market. Designing a technology radar is an established, yet mostly manual, process for visualizing recent technology trends.

The challenge of identifying and visualizing technologies is addressed by the project TechRad which uses machine learning to realize an autonomous technology scouting radar. One of its core functionalities is the identification of technologies in text documents. This is implemented via Natural Language Processing (NLP).

This paper aims to summarize the challenges and possible solutions for using entity recognition to identify relevant technologies in text documents. The authors present an early stage of implementation of the entity recognition model. This contains the selection of Transfer Learning as a suitable method, the creation of a dataset consisting of different data sources, as well as the applied model training process. Finally, the performance of the chosen method is benchmarked and evaluated in a series of tests.

Keywords

Machine Learning; Technology Management; Natural Language Processing; Entity Recognition

1. Introduction

This chapter aims to give a brief introduction to the problem at hand, describing the overarching research goal of the automated technology radar as well as providing motivation and context for the solution chosen in this work.

1.1 Challenge

Due to the constantly growing number of technologies available on the market, the number of devices using different digital technologies is growing exponentially [1]. This observation, alongside the fact that the time until a certain technology is known to a large number of users is diminishing [2], indicates that the frequency at which both companies and private users are exposed to new technologies is rising [3].

Organizations have to innovate in order to succeed and stay relevant in the market [4]. Mastering the process of finding technologies and managing innovations is a key success factor to ensure a company's market position [5]. Being unable to oversee the growing technology market endangers companies' long term strategic positions and ultimately their market position, and may even result in bankruptcy [3].

1.2 Solution

We are addressing the aforementioned issue by designing an automated technology scouting radar that gives an overlook over recent technological trends while keeping the research effort to a minimum. The software uses recent advancements in the field of Natural Language Processing (NLP), a sub-field of artificial intelligence (AI), to scout for information about technology in various sources. In a previous publication, the authors proposed an architecture for the automated radar [3].

The current paper focuses on a core functionality of the tool, namely identifying technologies in a written document. This functionality is realized using Named Entity Recognition (NER). The authors present the steps involved in the building and evaluation of the proposed Named Entity Recognition model.

1.3 Structure of the paper

Following the introduction in section 1, section 2 gives information about the state of the art and basic definitions. Section 3 summarizes the architecture of the automated radar to put the functionality of recognizing technologies into context. Section 4 describes the data sourcing and model building process in detail. After discussing the results in section 5, section 6 summarizes the insights. Section 7 concludes the paper by giving an outlook to future research in the area.

2. State of the art

The following paragraphs provide information about the state of the art and basic definitions to ensure a common understanding of the topics and terms used in the solution.

2.1 Technology management and visualization

A technology radar is a tool to summarize and visualize the results of a technology scouting process that organizes technologies in a circular diagram. The diagram is divided into sectors for structuring the content and delimiting the search areas, e.g., trends, technology fields, production technologies or product functions [5]. Figure 1 shows an example of a technology radar.

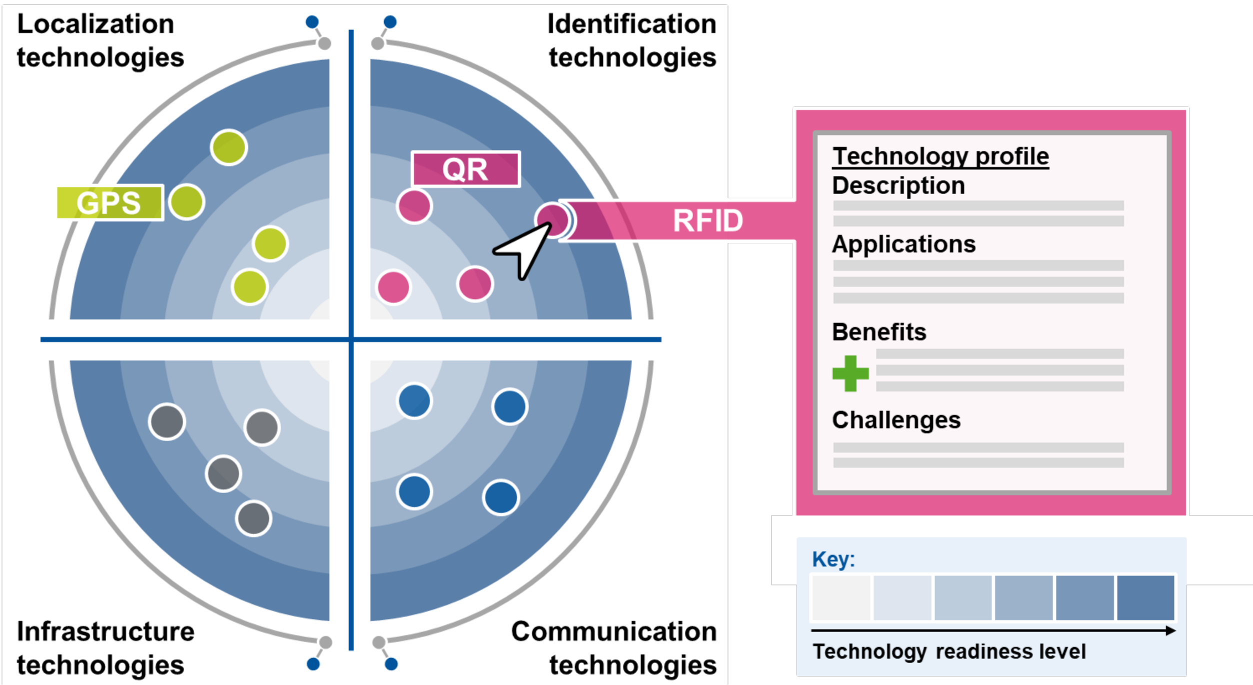


Figure 1: An example of a technology radar [6]

A temporal perspective is mapped on the axes, mostly the technology maturity, which indicates whether a search field or a specific technology is close to market readiness or in a research phase [5]. The underlying architecture proposed in our solution is explained in detail in section 3.

2.2 Data collection

Data collection is one of the main tasks when building AI models. Collecting data is not a trivial undertaking and demands a coherent approach. Especially in broadly scoped endeavours like the presented solution, the scientist encounters challenges like the lack of publicly available resources as well as copyrights [7]. The problem discussed in this paper demands a specialized corpus of data about emerging technologies. This is realized by using standardized API queries from scientific publication portals. The approaches and sources of data used are explained in detail in section 4.

2.3 Named Entity Recognition

A named entity is defined as a word or a phrase that clearly identifies one item from a set of other items that have similar attributes [8]. Named Entity Recognition (NER) is the problem of identifying sections of a text or certain words that mention named entities, and to subsequently classify them into predefined categories [9]. NER serves is a core functionality of many natural language applications such as translation, context sensitive answering and summarization [7]. In this paper, we present the use of NER in automatically identifying technology terms from a corpus using the advancements in deep learning techniques.

3. Architecture including identified functions

The previous work focused on the design process of a possible architecture for the radar [3], as it is a crucial preliminary step of the software and systems engineering process [10]. In this section, a brief summary of the process and results is given to ensure that the context in which Named Entity Recognition is used becomes clear (see also Figure 2).

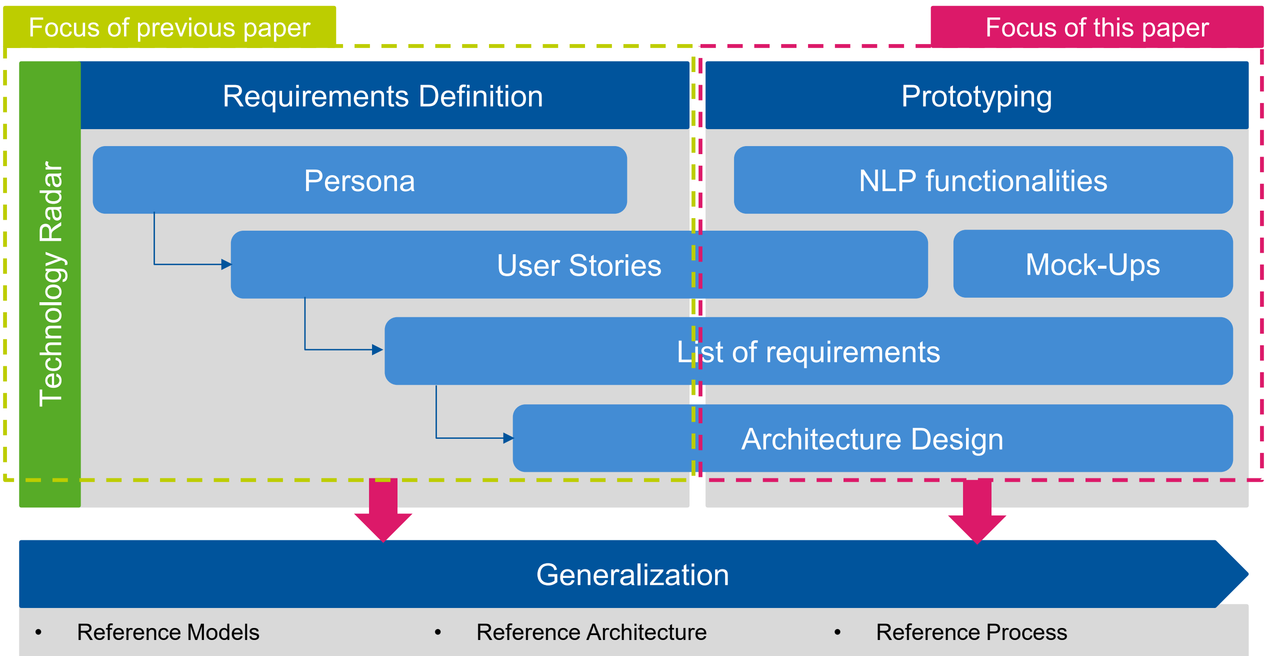


Figure 2: Overall project approach and focus of the paper [3]

3.1 Gross Functional Structure and Modules of TechRad

The automated technology radar is a solution consisting of four main steps. In data gathering, API queries and a web crawler are used to gather data from public sources, such as scientific publication portals, blogs and social media [3]. In the training phase presented in this paper, we used data exclusively from scientific publication portals and standardized social media APIs, planning on broadening the scope in further phases of the project. In the second step, data storage, the documents are allocated and organized for further processing. During the analysis the Full text documents will only be stored temporarily (cache). A document's metadata is stored for future reference. The third step is the analysis, in which the text files are checked for technologies using NER, the main focus of the paper. In a successive analysis step the technologies will be classified into technology readiness levels using different NLP methods. The fourth step is largely focused on the front-end and deals with user-friendly visualization of the extracted information. The design of the radar is presented in detail in [3]. A graphical presentation of the steps can be found in Figure 3.

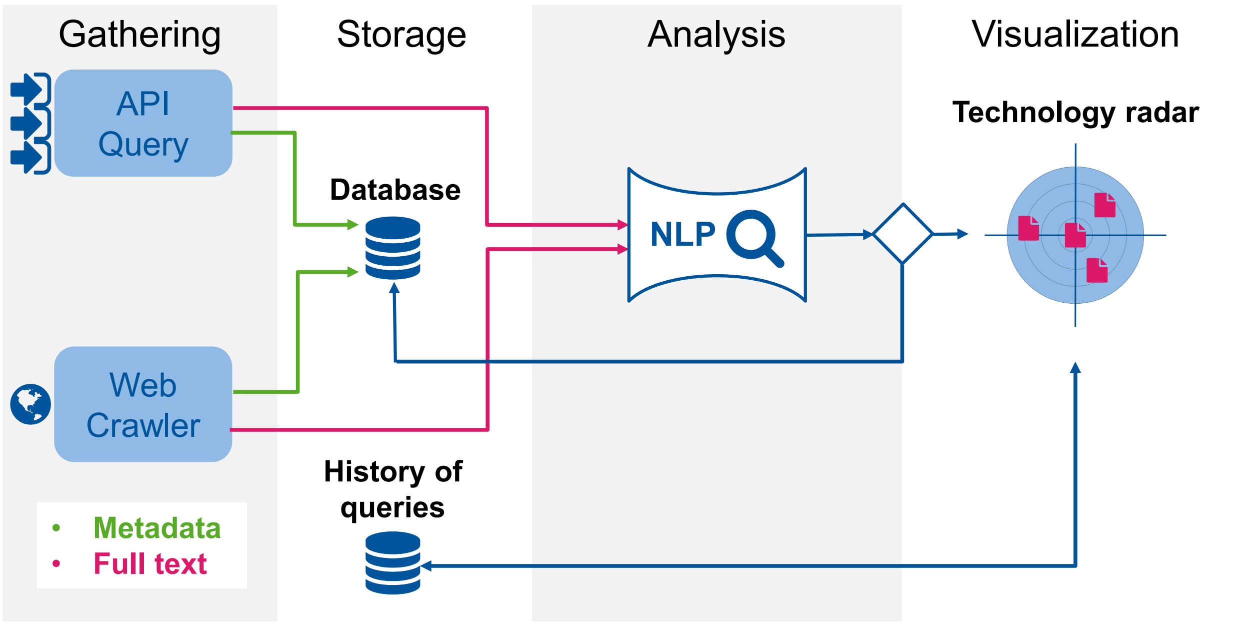


Figure 3: Architecture of the Radar [3]

4. Entity recognition for the identification of technologies

The following section describes the stages involved in building the Named Entity Recognition model. The success of the implementation depends on the quality of data sources, the pre-processing of data which is fed to the model and the hyperparameters used in the actual training procedure.

4.1 Data sources

With recent advances in machine learning, fostered by techniques such as deep learning, many tasks can be solved once a sufficiently large dataset is available for training. Nevertheless, human-annotated datasets are often expensive to produce, especially when the labels are used in high volume and frequency, as it is the case with word-level-annotation in NER [11].

Various sources were explored to build the necessary training data. As the application demands data rich in emerging technology terms, newly published scientific documents were considered as ideal sources. ArXiv [12] is a free distribution service and an open-access archive for scholarly articles in many emerging fields.

In addition to ArXiv, other major sources of rich scientific information like Google Scholar [13] and the Database and Logic Programming portal (DBLP) [14] were used.

To supplement the above-mentioned data sources, social media data were also collected through standardized API queries. A list of keywords such as Text Mining and Natural Language Processing were used as search terms through API queries in social media sites like Twitter. After the collection of sufficient quantity of data, pre-processing and training were performed.

4.2 Pre-processing

As an initial step, pre-processing of the data involves selecting chunks of texts from the collected data which are relevant for the application. The abstracts of the collected scientific papers were extracted and used as the training data. The next step involves preparing our data by annotating our text with "technology" tags. This is a labour-intensive task and is accomplished with the help of an open source text annotation tool for humans called "Doccano".

SpaCy is a free open-source library for Natural Language Processing in Python which features in-built NER [15]. Even though pre-built SpaCy models are good at NER extraction, they are not good enough for customized applications as the training data used is not specific to latest technologies. Therefore, the training data is manually labelled using Doccano.

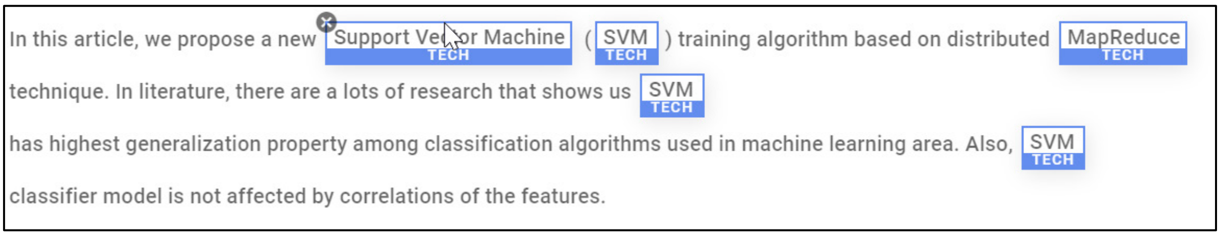


Figure 4: Training data annotation

Figure 4 shows a sample of the training data annotation. A technology entity is defined in the context of our application in the following way: if the entity is an algorithm, a commercial product, library or a framework, it is considered as a technology entity. Standard text corpus, performance evaluation metrics and broad umbrella terms were not considered as technology entities. Doccano provides a GUI for data annotation and the output is stored in a JSONL delimited file with each sentence along with the entity offsets. Figure 5 shows a sentence of the output JSONL file. The task of pre-processing also included procedures like un-latexing.

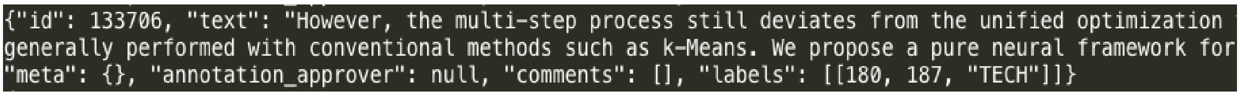

{"id": 133706, "text": "However, the multi-step process still deviates from the unified optimization
generally performed with conventional methods such as k-Means. We propose a pure neural framework for
"meta": {}, "annotation_approver": null, "comments": [], "labels": [[180, 187, "TECH"]]}

Figure 5: Sample JSONL file

4.3 Training procedure

The following paragraph describes the training procedure of the Named Entity Recognition algorithm. The SpaCy projects repository includes various project templates for different Natural Language Processing tasks, models, workflows and integrations. It allows to manage end-to-end SpaCy workflows for different use cases and domains. For this application, a project template for NER was selected and the pipeline was customized for better performance.

The output files from Doccano were converted into a SpaCy compatible format and stored, which was used as the training data in the pipeline. SpaCy uses a config file that contains all the model training components to train the model like component type, which is NER in this case, hardware accelerator selection and optimization goals. The algorithm was trained using ‘SciBERT’ as Transformer with dropout and ‘Adam.v1’ as the optimizer. Figure 6 shows an exemplary annotation performed by our algorithm on the test data.

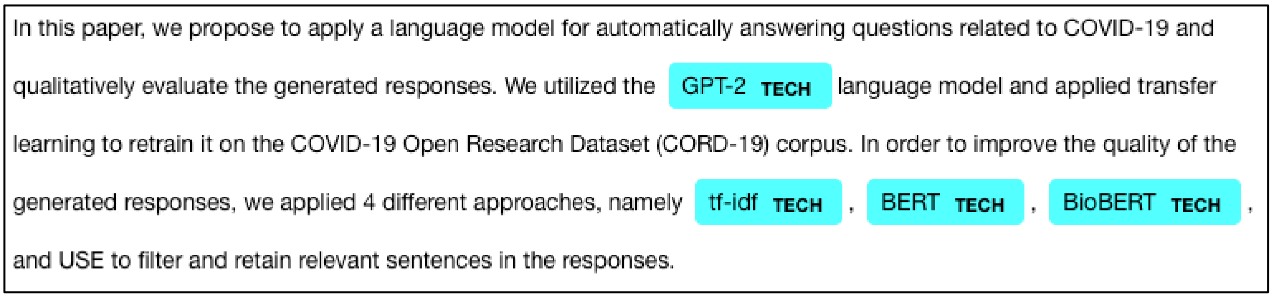


Figure 6: Annotation by trained NER algorithm

5. Results and Evaluation

The results, evaluation and benchmarks are presented in this section. Table 1 represents the evaluation of our algorithm on the test data based on a 70/30 test train split. The column ‘Epochs’ represents the number of epochs in training. The column ‘Iterations’ represents the number of iterations or steps in that particular epoch, ‘F-Score’ represents the F-Score, which is the harmonic mean of precision and recall. For a Named Entity Recognition task, the ideal ‘F-Score’ should be close to 1, which would represent a perfect model. The columns ‘Precision’ and ‘Recall’ represent the metrics precision and recall respectively. Precision denotes the percentage of predicted annotations that were correct, while recall denotes the percentage of reference annotations rightly recovered. Both these metrics should increase close to 100 for an ideal model. Our model was trained on a total of 1050 positive samples containing ‘TECH’ entities.

Table 1: Evaluation of the NER model

Epochs	Iterations	F-Score	Precision	Recall
4	4000	0.74	70.72	77.47
4	4200	0.73	77.68	68.49
4	4400	0.71	73.47	69.62
4	4600	0.71	65.37	78.61

5.1 Benchmarking in the general context

Transformer-based neural architectures are changing the field of NLP with an attention-based mechanism that outperforms convolutional or recurrent models [16]. Current NLP models are mostly based on deep neural networks which are characterized by great performance but are notoriously opaque in their prediction process [17]. Although different standard metrics have been proposed to standardize the evaluation, the authors stick to the common machine learning metrics like F-Score for evaluation, considering the ease of comparison between algorithms.

The F-Score, Precision and Recall shown in Table 1 are matching expectations of the research team and are close to similar efforts, where transformers have been used to solve NER problems on scientific corpora (see Table 2). These are comparable in so far as similar procedures have been used (fine-tuning a pre-trained model) and are tested on scientific corpora. However, the table is used to support the point that an F-Score of around 0.7 is acceptable with this training method applied to a scientific corpus; a direct performance comparison is not made, as the validation dataset would have to be the same, which is not desired at the current state of the prototype.

Table 2: Named Entity Recognition System evaluation[1]

Named Entity Recognition System	F-Score	Source
SpERT	0.73	[18]
RDANER	0.69	[19]
Cross-sentence	0.68	[20]

5.2 Critical Reflection

NER models are often trained based on formal documents and publications. Informal web documents that would be incorporated into the data basis by using web crawling or more social media sources usually contain noise, as well as incorrect and incomplete expressions. The performance of current NER systems generally decreases as informality increases in web documents [21]. This can be rectified using some post-processing, but could still pose a challenge in the further phases of development.

In general, deep learning techniques are data-hungry and the performance of these models increases with more data. Adding more training data could further improve the performance of our model, however the labelling process proved to be resource intensive. Another lever for better scores is hyperparameter optimization and tuning the config-parameters during the training process. For a prototype of the technology radar, an F-Score of about 0.7 is thought to be sufficient as it demonstrates technical viability and feeds the successive steps with relevant data to process and visualize. Still, the authors plan on increasing the precision of the model in further stages of the project.

6. Summary

In this paper, the authors present a prototype of an autonomous technology radar using NLP. The focus of the presented research lies on a core functionality problem that was solved using Named Entity Recognition. In the beginning, the need for automation in technology scouting is explained. The architecture of the autonomous radar is described, followed by the implementation and the evaluation of a prototype using NER to successfully identify technologies in text documents. The authors then evaluate the usefulness of the approach, which is deemed sufficient for the current state of the implementation.

7. Outlook

With the prototype showing promising results, the feasibility check for using Named Entity Recognition to develop an autonomous technology radar is performed successfully. The prototype can be used as a first step to build the final visualization of the results. Further research will focus on the steps to improve the accuracy of the algorithm. The architecture and the algorithm will be further evaluated with industry experts and potential users. Based on the feedback, the process will be fine-tuned.

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

[1] benchmarking conducted with https://paperswithcode.com in March 2021

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] Parasuraman, A., 2000. Technology Readiness Index (Tri). Journal of Serivce Research 2 (4), 307–320.

[2] Statista Research Development, 2018. Informationstechnologien: Entwicklung bis 50 Millionen Nutzer. https://de.statista.com/statistik/daten/studie/298515/umfrage/entwicklung-ausgewaehlter-informationstechnologien-bis-50-millionen-nutzer/. Accessed 21 January 2020.

[3] 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.

[4] Hoerlsberger, M., 2019. Innovation management in a digital world. JMTM 30 (8), 1117–1126.

[5] Schuh, G., Klappert, S., 2011. Technologiemanagement. Springer Berlin Heidelberg, Berlin, Heidelberg.

[6] Stroh, M.-F., Benning, J., 2020. Trends finden und überblicken für KMU – Teil 1. MM MaschinenMarkt 126 (8), 60–61.

[7] Zeroual, I., Lakhouaja, A., 2018. Data science in light of natural language processing: An overview. Procedia Computer Science 127, 82–91.

[8] R. Sharnagat, 2014. Named entity recognition: A literature survey: Center For Indian Language Technology.

[9] Nadeau, D., Sekine, S., 2007. Named Entities: Recognition, classification and use. LI 30 (1), 3–26.

[10] ISO, 2011. Systems and software engineering - Architecture description, 2011st ed., 48 pp.

[11] Menezes, D.S., Savarese, P., Milidiú, R.L., 2019. Building a Massive Corpus for Named Entity Recognition using Free Open Data Sources. http://arxiv.org/pdf/1908.05758v1.

[12] Cornell University. https://arxiv.org/. Accessed 17 May 2021.

[13] Alphabet. https://scholar.google.de/. Accessed 17 May 2021.

[14] Schloss Dagstuhl - Leibniz Center for Informatics. Schloss Dagstuhl - Leibniz Center for Informatics. https://dblp.org/. Accessed 17 May 2021.

[15] Industrial-Strength Natural Language Processing, 2021. https://spacy.io/.

[16] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need, 15 pp. http://arxiv.org/pdf/1706.03762v5.

[17] DeYoung, J., Jain, S., Rajani, N.F., Lehman, E., Xiong, C., Socher, R., Wallace, B.C., 2019. ERASER: A Benchmark to Evaluate Rationalized NLP Models. http://arxiv.org/pdf/1911.03429v2.

[18] Eberts, M., Ulges, A., 2019. Span-based Joint Entity and Relation Extraction with Transformer Pre-training. http://arxiv.org/pdf/1909.07755v3.

[19] Yu, H., Mao, X.-L., Chi, Z., Wei, W., Huang, H., 2020. A Robust and Domain-Adaptive Approach for Low-Resource Named Entity Recognition, 297–304.

[20] Zhong, Z., Chen, D., 2020. A Frustratingly Easy Approach for Entity and Relation Extraction. http://arxiv.org/pdf/2010.12812v2.

[21] Kim, M.H., Compton, P., 2012. Improving the Performance of a Named Entity Recognition System with Knowledge Acquisition, in: ten Teije, A., Völker, J., Handschuh, S., Stuckenschmidt, H., d'Aquin, M., Nikolov, A., Aussenac-Gilles, N., Hernandez, N. et al. (Ed.), Knowledge Engineering and Knowledge Management, vol. 7603. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 97–113.

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. He has had entrepreneurial success with multiple start-ups and spin-offs and is a member of the Presidium of the German Academy of Technical Sciences (acatech).

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 intelligent products. As head of the division, he is responsible for multifaceted consulting and research projects.

Max-Ferdinand Stroh (*1991) is a researcher at FIR at RWTH Aachen University since 2017 in the department Information Management. He is leading the group Information Technology Management at FIR. He is also deputy 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 – the main focus of his current research and publications.

Clinton Gnanaraj Charles (*1993) is a master student at RWTH Aachen University studying Data Analytics and Decision Science since 2019. He has a Bachelor's degree in Automobile Engineering, followed by 4 years of work experience in the automotive industry in India and United States. Fascinated by the potential of data and artificial intelligence, he decided to pursue further education in this field. He has been associated with the FIR as a research assistant since 2020.

2nd Conference on Production Systems and Logistics

Simulation-Based Evaluation Of The Hub-And-Spoke Concept To Support The Centrally Managed Supply Of Urban Factories

Patrick Reineke, Annika Sophie Geusch, Lennart Büth,
Mark Mennenga, Christoph Herrmann

Chair of Sustainable Manufacturing and Life Cycle Engineering, Institute of Machine Tools and Production Technology, Technische Universität Braunschweig, Langer Kamp 19b, 38106 Braunschweig, Germany

Abstract

The progressive urbanization and increasing decentralization of manufacturing lead to a growing need for the integration of manufacturing plants into urban areas. As a result, industrially induced traffic volumes and emissions are increasing in these areas. Therefore, it is important to manage urban logistics as effective as possible to reduce costs and environmental impacts. A promising approach to support the effective supply of urban factories can be seen in a centrally managed supply. Especially the hub-and-spoke concept is known to solve the problem of the last mile, e.g. within the parcel-industry. That is why most past research activities about hub-and-spoke concepts focus on deliveries to individuals or deliveries in the retail sector while only considering parcel deliveries. However, differing from the parcel industry the supply of urban factories requires especially just in time and just in sequence deliveries. Therefore, new approaches are required which help to analyse if and how a centrally managed supply of urban factories can support the reduction of traffic volumes and with this costs and environmental impacts. In this contribution we conduct a feasibility study on the effectiveness of a centrally managed supply of urban factories using a simulation approach. Based on the analysis of possible stakeholders and different hub concepts, we take a closer look on the application of the hub-and-spoke concept for urban logistics and provide a simulation based evaluation realized with the software Simulation of Urban Mobility (SUMO).

Keywords

Urban Factories; Logistics; Simulation-based evaluation; Emissions; Traffic Volume

1. Introduction

In the year 2018, more than half of the world´s population lived in urban areas, while forecasts predict further growth [1]. Because of that, urban factories are facing new challenges. In principle, there are two kinds of urban factories: On the one hand, there are factories, which originally were build outside of the city and, because of the urbanization, are now producing right next to residential areas. On the other hand, there are those factories, which are planned and build right in the city to use potentials of urban areas because of new production technologies and business models. [2] A decisive factor for the acceptance for both kinds of urban factories is the design of effective logistics flows within the city, which can help reducing the traffic volume and meeting emission limits. German cities frequently struggle with high pollutant emissions resulting from high traffic volume. Therefore, new approaches are needed to lower the traffic volume, especially the one resulting from industry-related logistics. One approach, which tackles the challenge of the "last mile" within the logistics of the parcel industry, is the hub-and-spoke concept [3–5]. In this approach, a centrally managed node (= hub) functions as a decoupling element between the decentrally organized flow of goods to the city boundaries and the centrally organized deliveries within the city. Even though the hub-and-spoke concept promises a high potential, it has not been evaluated thoroughly in the context of industry-related deliveries. Furthermore, production specific characteristics, e.g. just in time deliveries, have to be considered. That is why this work focusses on conceptualizing the centrally managed supply of urban factories and aims to conduct a feasibility study using a simulation-based approach. To this end, the paper is structured as follows: Section 2 shows the relevant state of research as a basis for the introduction of the methodology in section 3. Section 4 then illustrates the development process of the simulation-based approach, which is exemplarily applied in a case study in section 5. Finally, section 6 concludes this work and gives an outlook.

2. State of research

The importance of sustainable inner-city logistics has been emphasized by a wide variety of researchers, e.g. [6,7]. In literature, different approaches are proposed which aim at achieving a city-compatible and economical supply of goods while also considering the ecological needs of the environment. For example, Bernsmann and Vashtag [6] consider the creation of incentives instead of driving bans for freight traffic as an adequate approach and propose, among other things, electric mobility, night deliveries and small logistical locations in the heart of the city as possible solutions. Russo and Comi [8] investigate the relationships between city characteristics and the possible reduction of pollutant emissions and present a tool that can be used to identify promising measures to reduce pollutant emissions in a given city. Hesse [9] recommends traditional forms of planning as well as cooperation between companies or between municipal planning and companies as building blocks for a city logistics concept. Moreover, the use of various instruments is examined to determine ecological or urban quality standards for freight transport in new planning projects. Furthermore, an important and complex task is finding the best long-term location for a facility, which is why Kik et al. [10] introduce the Regional Facility Location and Development Planning Problem (RFLDP) and present a model that supports a company´s process of selecting a facility location. Gevaers et al. [11] use best practices and case studies to describe new concepts for solving inner-city delivery problems of the last mile. In addition to the delivery of mail to individuals, the transport of goods to companies and factories is also considered and technological innovations are highlighted. Göpfert [12] addresses numerous logistics visions and successful strategies for companies. The latest topics in future research for logistics, such as autonomous data-driven logistics, sustainable supply chains in the sense of the last mile and flexibility in complex supply networks, are dealt with. Besides the topic of the last mile, the necessity of data-driven approaches is also being examined in various papers: Ehmke offers a quantification of the value of dynamic and time-dependent information for improved route guidance in city logistics and connects different areas of traffic data acquisition with data mining and research techniques [13]. Furthermore, Montoya et al. [14] provide a quantitative assessment of the effects of collaborative approaches and mathematical models for optimizing inner-city freight traffic. There is also preliminary work for the simulation of transport orders for which free card and traffic data were used [15–17]. Taniguchi et al. [18] present recent trends and innovations in modelling city logistics with a focus on emissions, health care problems and logistics modelling for mega cities. Particularly for parcel deliveries the hub-and-spoke concept has been investigated and proven to be suitable to minimize the driven distance and time used for deliveries [19–22]. However, so far, little is known about its application to industry-related deliveries since there are no studies about the hub-and-spoke concept that consider production specific requirements, e.g. just in time deliveries. That is why this paper aims at filling exactly that research gap by investigating if a hub-and-spoke concept has any benefits in regards to driven distance and pollutant emissions in the context of urban factories.

3. Methodology

Based on the afore mentioned research gap an integrated approach for the evaluation of hub-and-spoke concepts for industry-related deliveries is proposed and exemplarily applied. The following research question is addressed: “How can the industry-related traffic volumes and the resulting pollutant emissions in cities be reduced using a hub-and-spoke concept?” The air pollution not only depends on emission-free drive concepts, but also on traffic disruption, e.g. congestions. Therefore the traffic management is very important to plan the traffic as effective as possible [23,24]. To do that, new forms of regulation for transportation planning processes and a broader database are needed [25]. The planning processes have to be aligned with each other, which is why a centrally managed organization makes sense. Accordingly, it can be hypothesized that a centrally managed supply of urban factories by using the hub-and-spoke concept may also reduce the traffic volume and resulting pollutant emissions. To prove this hypothesis, a simulation-based feasibility study is conducted. Therefore, first requirements for this concept are investigated by using a stakeholder-analysis based on a literature review. Afterwards the simulation approach is developed. It is based on the investigated requirements and considers the selection of different delivery car types that are usable in inner cities, just in time deliveries, a methodology to consolidate deliveries and different average speed of individual car traffic in inner cities dependent on the daytime. Finally, the simulation is used in a case study in the city of Hamburg. For the simulation the Software SUMO is used, which is a microscopic, free access traffic simulation that enables loading open source card data into it [26]. The simulation approach is closely linked to the development of the requirements of the stakeholders and a holistic evaluation system as well as the conception of distribution centres, means of transport planning and fleet combination. In this way, a comprehensive evaluation of sustainable inner-city logistics concepts can be guaranteed. The approach further offers potential for scenario-based analysis and redesign of inner-city logistics concepts.

4. Development of the simulation study

As a foundation of the simulation-based approach, a stakeholder analysis is implemented (see Figure 1).

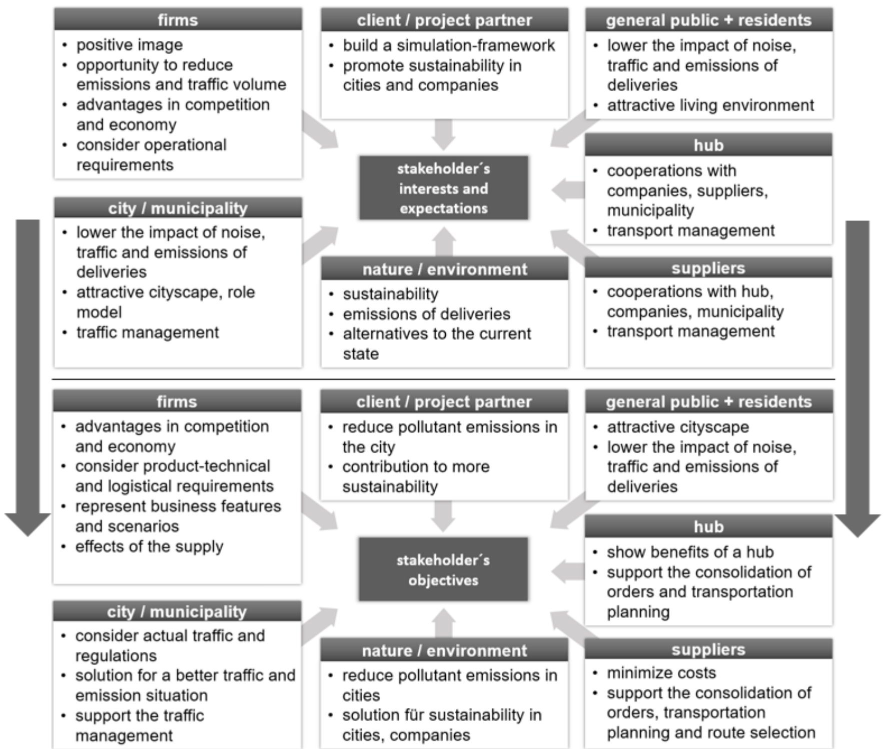


Figure 1: Stakeholder´s interests and expectations (on top), stakeholder´s objectives (bottom), derived based on [27,28]

The stakeholder analysis is based on a qualitative literature review and the approach described by [27,28]. It serves the derivation of the interests and expectations of all relevant stakeholders and consequently of their objectives as a basis for the simulation. The stakeholder analysis comprises firms, clients and project partners / clients, general public and residents, the hub itself, suppliers, the city or municipality as well as the nature and environment. Thereby, firms and municipality can be seen as the most important stakeholders in regards to the implementation of a simulation-based evaluation of the hub-and-spoke concept, since those are the possible end users of the simulation and the subsequent evaluation.

Based on the stakeholder objectives, which are derived from the stakeholder´s interests and expectations, the technical and organizational requirements of a hub, the requirements on the formation of transport orders for industry-related delivery as well as the requirements for the simulation study are formulated as illustrated in Figure 2. The technical and organizational requirements of a hub are derived from the stakeholder objectives and serve as a basis for the formation of transport orders for industry-related delivery, which should consider the shown requirements. Also based on the findings of the stakeholder-analysis, it is possible to generate the requirements for the simulation study that are shown as well.

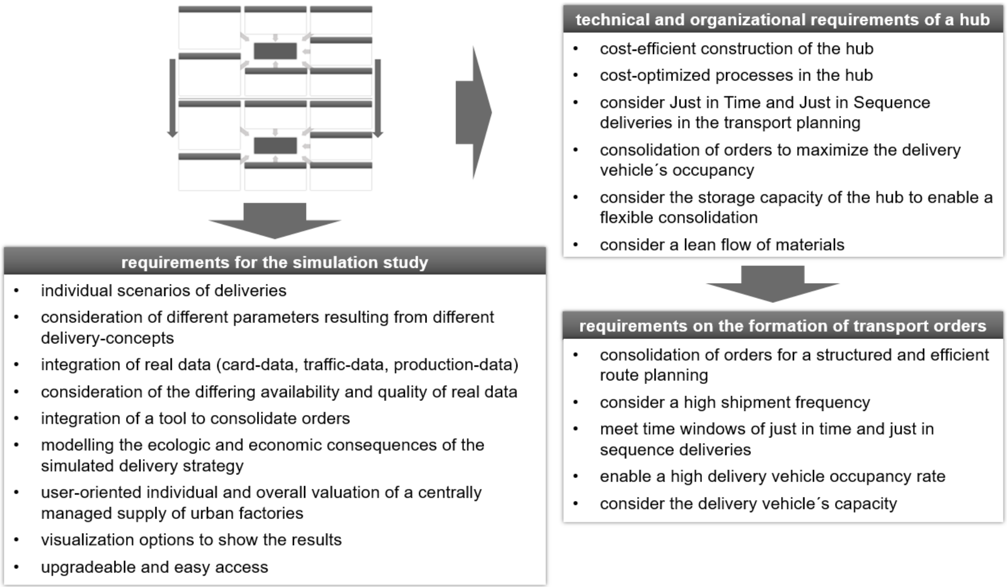


Figure 2: Requirements for the simulation study, technical and organizational requirements of a hub and requirements on the formation of transport orders

With the knowledge about requirements, the simulation approach can be build. It consists of five different modules: 1) the setup-module, 2) the firm-module, 3) the activity-module, 4) the simulation-module and 5) the visualization-module. The *setup-module* is used to build the environment within the simulation, e.g. streets from real card-data, traffic regulations, speed limits and locations of factories. The card-data can be downloaded from OpenStreetMap (OSM), which is an open system that can be edited and shared by users free of charge. In regards to OSM it has to be taken into account that the data is not suitable to use for use-cases where a high level of detail is needed since specific data is missing sometimes, e.g. broad width or acceleration lanes [29]. Furthermore, a few errors of the integration process of the OSM-data into SUMO are existent: Traffic lights sometimes do not function as they should and the traffic regulations are faulty when there are too many vehicles in the simulation. In addition, there is only data about speed limits, but not about traffic demand. In reality, the average driven speed is lower than the speed limit, especially in the inner city [30]. The actual speed of vehicles is affected by many factors, such as seasonal characteristics, the daytime or weather conditions. This information often gets integrated into a simulation by using induction loop measurements. To counteract the described factors and to make the definition process of the simulation easier, average speeds can be used in dependence of the daytime. Figure 3 shows these average speeds. Hence, the most congested times are between six and eight am and between four and five pm.

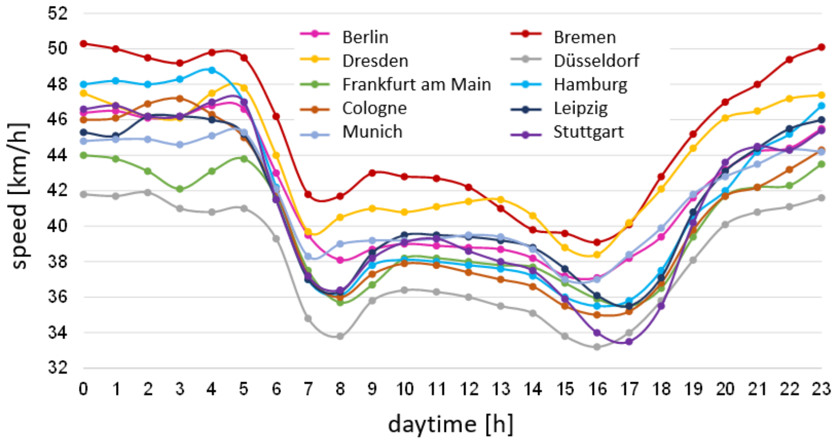


Figure 3: Average speed of individual car traffic in selected german cities (following [31])

The *firm-module* contains data about the simulated firms in an excel sheet, e.g. deliveries per day, weight of the deliveries, number of just in time deliveries. With this data, a python script generates a list of orders. This list is then used in the *activity-module* to generate a file in XML-format, which is used in SUMO to simulate the individual tours. For that, starting points and departure times of the vehicles are generated. If a hub should be considered, other restrictions apply: the time difference between two orders that may get consolidated to one is limited to avoid high stocks within the hub. Furthermore, the total weight of consolidated orders must not exceed the limit of the used delivery vehicle and the total route length as well as the number of stops is limited. While meeting these restrictions the best combination of orders to consolidate is chosen by following this priority:

1. The delivery vehicle should be as small as possible while being as fully laden as possible.
2. The orders that get combined preferably should have the same destination.
3. The occupancy of the delivery vehicle should be as high as possible.
4. The time difference between orders that may get combined should be as small as possible.

In addition, the orders get chronological ordered in a way that the distance is as small as possible while just in time deliveries will always be delivered first and are never combined with another just in time order. Moreover, the driven routes are always ending at the last destination, so the tours are not circular. The *simulation-module* consists of a python-script that starts the configuration-file of SUMO. This enables to change parameters during the running simulation, e.g. changing the average speed of vehicles on the simulated streets. Finally, the *visualization-module* tracks the results of the simulation and converts them into an excel sheet which shows data about departure and arrival times, duration of the tours, length of the routes and the sum of emissions and fuel consumption.

5. Case study: Hamburg

After building the simulation study, it is applied in an exemplary case study onto the city of Hamburg. Therefore, three possible classes of delivery vehicles are considered, which are allowed to drive in inner cities. These are (permissible total weight in brackets): Light duty-vehicles (< 3,5 t), light trucks (≥ 3,5 t and < 7,5 t) and medium-weight trucks (≥ 7,5 t and < 18 t) [32,33]. Figure 4 shows the selected section of Hamburg and the locations of plants and simulated hubs. This section is chosen because it includes a part of the inner city with its bigger access roads, e.g. highways, and therefore deliveries to the hubs are possible and the important main roads of the inner city are mapped. The locations of the plants are partly based on real locations of plants and partly based on the assumption that more firms will move into the city because of the urbanization. The three hubs, which are located near highway exits, represent three of four examined scenarios.

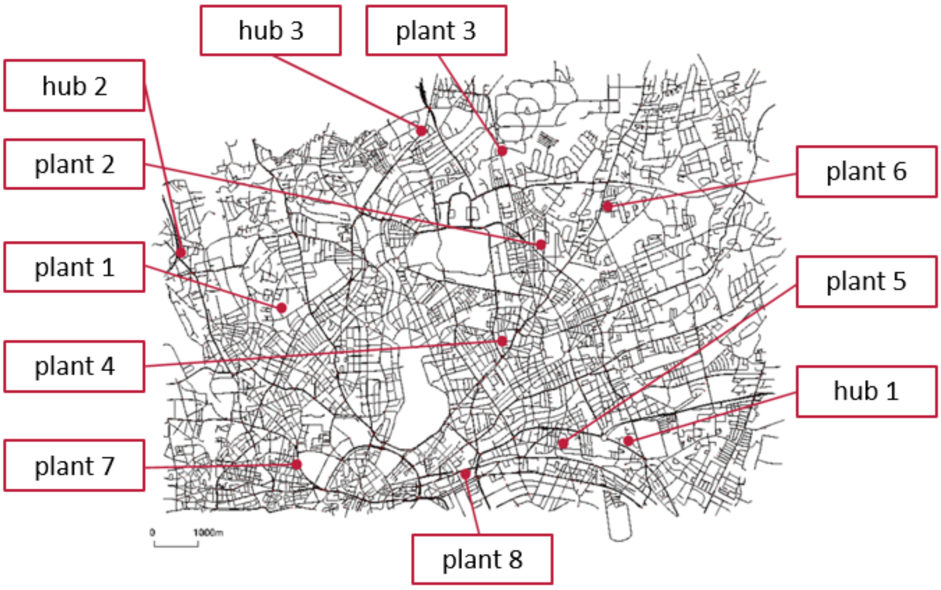


Figure 4: Selected section of Hamburg

These locations enable the decentrally organized supply of the hubs. The other scenario - or base-scenario - includes no hub, which means that the deliveries will start on federal roads and highways around the border of the shown section.

Table 1: Characteristics of plants

Name	Total number of orders simulation´s duration (1 day)	Share of JIT / JIS [%]	Time window JIT / JIS [s]	Minimum weight [t]	Maximum weight [t]
Plant 1	250	20	3600	0.5	9
Plant 2	250	10	3600	0.5	9
Plant 3	124	-	-	0.5	9
Plant 4	118	7	3600	1.2	9
Plant 5	80	-	-	0.5	9
Plant 6	50	-	-	0.5	9
Plant 7	20	-	-	0.5	9
Plant 8	8	-	-	0.5	9

As data for the firm-module, Table 1 serves as input. This information is based on (anonymized) company data. To measure the distance between plants, a preliminary study is conducted in which single delivery cars, whose route is tracked, drive the distances between the shown plants and hubs. The findings are then used in the real simulation to combine orders with a minimum route duration and length. Therefore, the following simplifications are made:

- The generated orders are all between six am and ten pm to avoid a possible ban on night driving.
- The flow of traffic is regulated by the described average speed.
- While consolidating, the volume of goods is not considered, only the weight.
- There is no differentiation between product types so that orders can get combined more freely.

Based on this, the simulation results can be generated. First, Table 2 shows the total number of deliveries and the total sum of driven kilometres. The figures in brackets are the difference to the base-scenario. In Scenario 3, the combination of driven kilometres and total number of deliveries is the lowest, but overall the scenarios that use a hub have a lower number of deliveries and driven kilometres as expected.

Table 2: Comparison of the scenarios

Parameter	Base-scenario	Scenario 1	Scenario 2	Scenario 3
Total number of deliveries (vehicles)	900	557 [-38.11%]	560 [-37.78%]	560 [-37.78%]
Total length of driven tours [km]	7476.63	4385.98 [-41.34%]	5290.99 [-29.23%]	4105.27 [-45.09%]
Average length of a tour [km]	8.31	7.87 [-5.29%]	9.45 [+13.72%]	7.33 [-11.79%]

Table 3 shows that the occupancy of the delivery vehicles is higher while using the hub-and-spoke concept.

Table 3: Selected vehicle types and their occupation

Parameter	Base-scenario	Scenario 1	Scenario 2	Scenario 3
Light duty vehicle				
number [vehicles]	64	26	26	26
occupancy [%]	78.52	94.23	94.23	96.35
Light trucks				
number [vehicles]	329	64	65	66
occupancy [%]	66.16	97.85	97.73	97.66
Medium-weight trucks				
number [vehicles]	507	467	469	468
occupancy [%]	71.01	92.58	92.10	91.39

Table 4 illustrates that the fuel consumption as well as the pollutant emissions are lower in the scenarios that use a hub. To estimate the cost of the driven distance a price of 1.30 Euro per litre diesel is assumed.

Table 4: Summary of the emission results

Parameter	Base-scenario	Scenario 1	Scenario 2	Scenario 3
Fuel consumption [l]	1729.87	1244.44 [-28.06%]	1489.87 [-13.87%]	1179.08 [-31.84%]
Fuel costs [€]	2248.83	1617.77 [-28.06%]	1936.83 [-13.87%]	1532.80 [-31.84%]
CO_2 [kg]	4338.95	3121.38 [-28.06%]	3734.95 [-13.92%]	2957.43 [-31.84%]
CO [kg]	8.81	6.27 [-28.83%]	7.40 [-16.00%]	6.02 [-31.67%]
HC [kg]	2.49	1.76 [-29.32%]	2.04 [-18.07%]	1.71 [-31.33%]
NO_x [kg]	41.23	29.47 [-28.52%]	35.08 [-14.92%]	28.05 [-31.97]
PM_x [kg]	1.01	0.69 [-31.68%]	0.81 [-19.80%]	0.66 [-34.65%]

The shown pollutant emissions are based on the emission classes, which are integrated into SUMO. These are derived from the handbook of emission factors of road traffic ("Handbuch für Emissionsfaktoren des

Straßenverkehrs" – HBEFA) while only direct emissions are considered [34]. The light duty vehicles are in the emission class HBEFA2/P_14_5, the light trucks in HBEFA2//HDV_12_2 and the medium-weight trucks in HBEFA2//HDV_12_12. The figures in the brackets are the difference to the base-scenario. When comparing Table 2 to Table 4 it is apparent that the emissions are not reduced to the same degree as the total driven kilometres are. This derives from the higher percentual amount of bigger vehicles in the scenarios that use a hub (see Table 3). In addition, all just in time deliveries in the simulation met their time windows while using a hub.

All in all the results of the case study show that the hub-and-spoke concept can reduce the total number of tours that take place in the inner city. Depending on the location of the hub, the reduction is up to 38%. Furthermore, the total driven distance can be reduced by 29% to 45%. The occupancy of the delivery vehicle can be improved from 66% to 78% in the base-scenario to 91% to 98% depending on the vehicle class and the location of the hub. In addition, the pollutant emissions of CO_2 are reduced by 14% to 32%. These results indicate that the hub-and-spoke concept in the context of urban production can lead to a higher quality of life by reducing traffic volume and pollutant emissions.

6. Conclusion and outlook

The progressive urbanization and increasing decentralization of manufacturing lead to a growing need for the integration of manufacturing plants into urban areas. As a result, industrially induced traffic volumes and emissions are increasing in these areas. Therefore, it is important to manage urban logistics as effective as possible to reduce costs and environmental impacts. This paper investigates the potential of the hub-and-spoke concept in the context of urban production by providing an integrated evaluation methodology and applying it to exemplarily compare different scenarios in a simulation of a selected area of Hamburg. The results show that there seems to be a high potential in regards to a possible reduction of pollutant emissions and traffic volume. For this, future work is necessary which should include a specific setup, e.g. specific firms and a city that offers real time traffic data. In addition, the related costs of hub-and-spoke concepts need to be investigated in more detail. Furthermore, the specific firm data could show restrictions and limitations of the concept to industrial practice, which then can get investigated further (e.g. storage capacities in the hub). Beyond that, new forms of mobility should get examined, e.g. load bicycles or electrical trucks.

Acknowledgements

This Paper is a shortened and translated summary of the final report of the project "Konzeption und Machbarkeitsuntersuchung einer zentral gesteuerten Anlieferung von urbanen Fabriken - ZAUF" (funding code: VB18F1029A), which was supported by the innovation programme mFUND of the german Federal Ministry of Transport and Digital Infrastructure.

References

[1] United Nations, Department of Economic and Social Affairs, 2019. World Urbanization Prospects: The 2018 Revision. United Nations, New York, 123 pp.

[2] Juraschek, M., Thiede, S., Herrmann, C., 2018. Urbane Produktion: Potenziale und Herausforderungen der Produktion in Städten, in: Corsten, H., Gössinger, R., Spengler, T.S. (Eds.), Handbuch Produktions- und Logistikmanagement in Wertschöpfungsnetzwerken. Walter de Gruyter GmbH, Berlin/Boston, pp. 1113–1133.

[3] Bode, W., 2016. Neue City-Logistik-Konzepte und -Techniken für mehr Nachhaltigkeit per City-GVZ und eStore, in: Deckert, C. (Ed.), CSR und Logistik. Spannungsfelder Green Logistics und City-Logistik. Springer Gabler, Berlin, Heidelberg, pp. 281–291.

[4] Bretzke, W.-R., 2015. Logistische Netzwerke. Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Gudehus, T., 2012. Logistik 2: Netzwerke, Systeme und Lieferketten, Studienausg. der 4. Auflage ed. Springer Vieweg, Berlin, 1 Online-Ressource.

[6] Bernsmann, A., Vastag, A. (Eds.), 2018. Urbane Logistik: Schnell, stadtverträglich und wirtschaftlich, 1. Auflage ed. huss, München, 160 pp.

[7] Strale, M., 2019. Sustainable urban logistics: What are we talking about? Transportation Research Part A 2019 (130), 745–751.

[8] Russo, F., Comi, A., 2012. City Characteristics and Urban Goods Movements: A Way to Environmental Transportation System in a Sustainable City. Procedia - Social and Behavioral Sciences 39, 61–73.

[9] Hesse, M., 1998. Wirtschaftsverkehr, Stadtentwicklung und politische Regulierung: Zum Strukturwandel in der Distributionslogistik und seinen Konsequenzen für die Stadtplanung. Deutsches. Institut. für Urbanistik, Berlin, 368 pp.

[10] Kik, D., Wichmann, M.G., Johannes, C., Spengler, T.S., 2019. An Integrated Multi-criteria Approach for the Regional Facility Location and Development Planning Problem, in: Bierwirth, Merkle (Eds.), Logistics Management. Springer International Publishing, Cham.

[11] Gevaers, R., van de Voorde, E., Vanelslander, T., 2009. Characteristics of Innovations in Last-Mile Logistics - Using Best Practices, Case Studies and Making the Link with Green and Sustainable Logistics. European Transport Conference, 2009 Proceedings, 1–21.

[12] Göpfert, I., 2016. Logistik der Zukunft - Logistics for the Future, 7. Aufl. 2016 ed. Springer Fachmedien Wiesbaden; Imprint: Springer Gabler, Wiesbaden, 1 Online-Ressource (XX, 499 S. 150 Abb. in Farbe).

[13] Ehmke, J.F., 2012. Integration of Information and Optimization Models for Routing in City Logistics. Springer US, Boston MA.

[14] Montoya-Torres, J.R., Muñoz-Villamizar, A., Vega-Mejía, C.A., 2016. On the impact of collaborative strategies for goods delivery in city logistics. Production Planning & Control 27 (6), 443–455.

[15] Barcel, J., Codina, E., Casas, J., Ferrer, J.L., Garca, D., 2005. Microscopic traffic simulation: A tool for the design, analysis and evaluation of intelligent transport systems. Journal of Intelligent and Robotic Systems 41 (2-3), 173–203.

[16] Hunt, J.D., Stefan, K.J., 2007. Tour-based microsimulation of urban commercial movements. Transportation Research Part B: Methodological 41 (9), 981–1013.

[17] Praticò, F.G., Vaiana, R., Gallelli, V., 2012. Transport and traffic management by micro simulation models: operational use and performance of roundabouts, in: Urban Transport XVIII. Urban Transport 2012, A Coruna, Spain. 15.05.2012 - 17.05.2012. WIT PressSouthampton, UK, pp. 383–394.

[18] Taniguchi, E., Thompson, R.G., Yamada, T., 2014. Recent Trends and Innovations in Modelling City Logistics. Procedia - Social and Behavioral Sciences 125, 4–14.

[19] Hu, L., Zhu, J.X., Wang, Y., Lee, L.H., 2018. Joint design of fleet size, hub locations, and hub capacities for third-party logistics networks with road congestion constraints. Transportation Research Part E: Logistics and Transportation Review 118, 568–588.

[20] Lin, C.-C., Chen, S.-H., 2004. The hierarchical network design problem for time-definite express common carriers. Transportation Research Part B: Methodological 38 (3), 271–283.

[21] Sim, T., Lowe, T.J., Thomas, B.W., 2009. The stochastic -hub center problem with service-level constraints. Computers & Operations Research 36 (12), 3166–3177.

[22] Zäpfel, G., Wasner, M., 2002. Planning and optimization of hub-and-spoke transportation networks of cooperative third-party logistics providers. International Journal of Production Economics 78 (2), 207–220.

[23] Lyons, T.J., Kenworthy, J.R., Newman, P., 1990. Urban structure and air pollution. Atmospheric Environment. Part B. Urban Atmosphere 24 (1), 43–48.

[24] Umweltbundesamt, 2020. Umweltbelastungen durch Verkehr. https://www.umweltbundesamt.de/daten/verkehr/umweltbelastungen-durch-verkehr#undefined. Accessed 2 March 2021.

[25] Batty, M., Axhausen, K.W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., Ouzounis, G., Portugali, Y., 2012. Smart cities of the future. Eur. Phys. J. Spec. Top. 214 (1), 481–518.

[26] Behrisch, M., Bieker, L., Erdmann, J., Krajzewicz, D., 2011. SUMO - Simulation of Urban MObility - An Overview. Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation. ThinkMind.

[27] Krips, D., 2017. Stakeholdermanagement - Kurzanleitung, 2. Auflage ed. Springer Vieweg, Berlin Heidelberg.

[28] Weiand, A., 2012. Stakeholder-Analyse - Interessengruppen identifizieren und einbinden. Zeitschrift Führung + Organisation 2012 (81), 134–137.

[29] Dallmeyer, J., 2014. Simulation des Straßenverkehrs in der Großstadt: Das Mit- und Gegeneinander verschiedener Verkehrsteilnehmertypen. Teilw. zugl.: Frankfurt, Univ., Diss., 2013 u.d.T.: Akteursorientierte multimodale Straßenverkehrssimulation. Springer Vieweg, Wiesbaden, 274 pp.

[30] Statista, 2008. Europäische Städte mit der niedrigsten Durchschnittsgeschwindigkeit im Straßenverkehr. https://de.statista.com/statistik/daten/studie/37200/umfrage/durchschnittsgeschwindigkeit-in-den-15-groessten-staedten-der-welt-2009/. Accessed 2 March 2021.

[31] BVL.digital/HERE Technologies, 2020. Accelerating Urban Logistics: Eine gemeinsame Studie von BVL.digital und HERE Technologies. https://go.engage.here.com/Accelerating-Urban-Logistics.html. Accessed 2 March 2021.

[32] Gabler, M., Schröder, S., Friedrich, H., Liedtke, G., 2013. Generierung der Nachfragestrukturen für die mikroskopische Simulation des städtischen Distributionsverkehrs im Lebensmittelhandel, in: Clausen, U., Thaller, C. (Eds.), Wirtschaftsverkehr 2013. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 32–48.

[33] Lenz, B., 2010. Shell LKW-Studie: Fakten, Trends und Perspektiven im Straßengüterverkehr bis 2030. Deutsches Zentrum für Luft- und Raumfahrt. http://www.shell.de/promos/media/shell-truck-study/_jcr_content.stream/1455802359533/6cae81224f158864b0b6ec3a048054ec49a85cb4159dbd0e1f41ed926bc27aff/publications-2010truckstudyfull.pdf.

[34] Deutsches Zentrum für Luft- und Raumfahrt, 2021. Emissions. https://sumo.dlr.de/docs/Models/Emissions.html. Accessed 20 February 2021.

Biography

Patrick Reineke (*1991) is currently a research assistant in the Institute of Machine Tools and Production Technology at the Technische Universität Braunschweig, Germany.

Annika Sophie Geusch (*1995) is a masters student in machine engineering at the Technische Universität Braunschweig.

Lennart Büth (*1988) is currently a research engineer and team lead for urban production in the Institute of Machine Tools and Production Technology at the Technische Universität Braunschweig, Germany.

Dr.-Ing. Mark Mennenga (*1983) is deputy head of the Chair of Sustainable Manufacturing and Life Cycle Engineering at the Institute of Machine Tools and Production Technology, Technische Universität Braunschweig. As research group leader "systems of system engineering" his main research areas are methods and tools for the engineering of sustainable systems, as well as of data-based business models and industrial product and production service systems.

Prof. Dr.-Ing. Christoph Herrmann (*1970) is university professor for Sustainable Manufacturing & Life Cycle Engineering and co-director of the Institute of Machine Tools and Production Technology, Technische Universität Braunschweig as well as director of the Fraunhofer Institute for Surface Engineering and Thin Films IST since November 2018.

Conference On Production Systems And Logistics
CPSL 2021

2nd Conference on Production Systems and Logistics

Development of a Practical Orientation Guide with Industrie 4.0 Use Cases for Industrial Manufacturers

Simon Schumacher[1], Bastian Pokorni[2], Roland Hall[1], Andreas Bildstein[1],Moritz Hämmerle[2]

[1]Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstr. 12, 70569 Stuttgart, Germany
[2]Fraunhofer Institute for Industrial Engineering IAO, Nobelstr. 12, 70569 Stuttgart, Germany

Abstract

Apart from innovators and early adopters, industrial manufacturers are still looking for adequate guidance for the orientation, prioritization, and development of Industrie 4.0 initiatives. Case studies of best practices have proven beneficial for the identification of use cases in manufacturing companies. Therefore, the aim of this paper is to develop an orientation guide for Industrie 4.0 with best practice use cases structured into thematic categories of production system areas. The orientation guide gathers relevant use cases for the adaption in industrial implementations. As a first practical application, the guide will be rolled out in an international study to primarily support regional manufacturing companies from Germany in the development of own Industrie 4.0 use cases. However, the guide is universally applicable to any manufacturing company and provides useful insights for the implementation of Industrie 4.0 use cases globally.

Keywords

Industrie 4.0; Production System; Use Cases; Production Strategy.

1. Introduction

The economy is crossing the threshold of the fourth industrial revolution and digitalization is already influencing the daily lives, economy and workplaces. A wide variety of areas of everyday life are already heavily influenced by information and communications technologies (ICT). Mobile internet, smartphones, cloud and distributed collaboration have led to major changes in the recent years. The widespread entry of ICT into industry is already in full progress, and there is no sign of this development slowing down. It is leading to a reshaping of entire value chains through the multi-layered use of the Internet of Things, mobile networking and flexible robotics. [1–3]

The aim of Industrie 4.0 in Germany is to strategically establish the country as a lead market, including leading technology providers. This offers medium-sized, export-oriented companies in particular from the mechanical engineering sector the potential to position themselves worldwide as pioneers of new, innovative solutions. Driven by the internet, the real and virtual worlds are continuing to grow together into an internet of things, services and people. Industrie 4.0 is an important research stream for the management of growing complexity in socio-technical production systems. [4]

Industry in the region of Baden-Württemberg has already taken up the topic. In particular, large companies are driving forward the digital transformation of their value chains and their range of services. Use cases for artificial intelligence, innovative industrial implementation solutions, new product-service combinations and business models are emerging and some are already in use. [5] In this context, overarching topics such as the creation of standards, IT safety and security, legal framework conditions, labor market trends, qualification measures and the effects on work systems design and work organization are currently being discussed and shaped.

Many companies still find it difficult to relate their own digitized production processes to typical Industrie 4.0 categories. [6] As a practical orientation support for regional companies, which do not yet know exactly whether their implemented solutions are feasible, an orientation guide for Industrie 4.0 best practices is meant to be created.

For the orientation, structuring and localization of these implementation solutions, a systematic guidance is to be developed on the basis of nationally and internationally collected Industrie 4.0 use cases. For this purpose, the *Framework for Cognitive Production Work 4.0* from [7] will be adapted to the practical needs of manufacturing companies in search of innovative implementation cases. The aim is to empower companies to quickly classify their application cases for Industrie 4.0 and at the same time identify gaps in their digitalization portfolio. The orientation guide forms the basis of a comparative study on Industrie 4.0 success factors and best practices.

This paper is structured as follows. Chapter 2 gives an overview of the relevant terminology and existing related works. In chapter 3 the development of the orientation guide is explained with regard to the thematic structure and categories, the application process, and its current limitations and potentials. Chapter 4 provides an outlook on the upcoming research activities and practical application of the orientation guide. The concluding remarks are gathered in chapter 5.

2. Related works

In this chapter the relevant related works for this paper are presented. For this, a definition of the relevant terminology (*Industrie 4.0, digital transformation, use case, Cognitive Production Work 4.0*) is provided in section 2.1. In order to clarify the underlying research framework for the following developments, the specific concepts are briefly explained in section 2.2.

2.1 Terminology

Industrie 4.0 describes the intelligent networking of people, machines and processes in industrial production with the help of information and communications technology (ICT). In this context, Industrie 4.0 is attributed with the change potential of a fourth industrial revolution. The term *digital transformation* in production refers to the implementation process of Industrie 4.0 and consists of the four phases of digitization, virtualization, networking and autonomization. [2,3,8,9] In a recent analysis of influencing factors of the digital transformation on supply chain complexity dimensions, [10] define the most important drivers and enabling technologies for the digital transformation per phase, which are displayed in Figure 1.

Industrie 4.0 technologies are often introduced in a strategic road mapping process with the help of concrete use cases. [11] A *use case* describes an exemplary application of a technology, method or tool. A use case contains concrete technical solutions to problems and also how these solutions can be practically implemented. In the operational context, use cases can be linked to involved actors, business processes and relevant key figures. [12] Use cases can be used to address individual lean and digitalization maturity levels of different plants in a manufacturing enterprise. This enables the benefits of the application to be communicated using practical examples. In addition, use cases can include the estimation and initial calculation of financial and other measurable potentials.

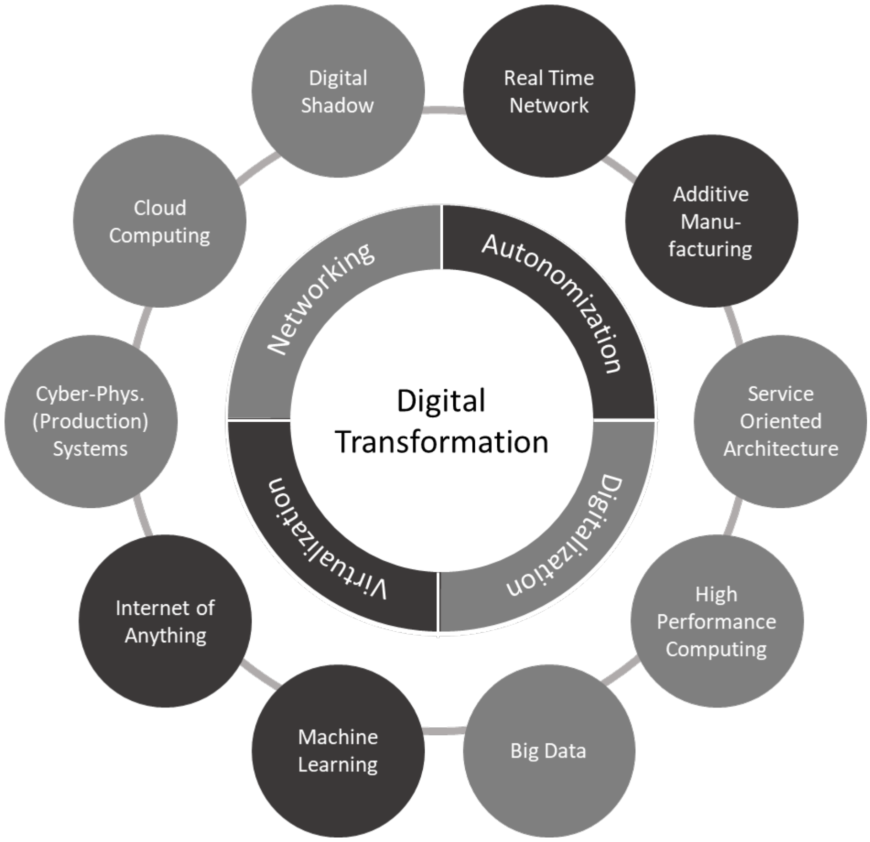


Figure 1: Drivers and enabling technologies for the digital transformation [[10] based on [13]]

Cognitive Production Work 4.0 is defined as follows. Production work describes work within manufacturing processes and will undergo significant change as a result of the use of Industrie 4.0. In terms of growing content, the increasing integration of ICT within work processes will result in a shift from executive activities to more controlling and regulating tasks. At the same time, flexibility requirements in terms of content, time and place will increase. *Cognitive Production Work 4.0* additionally describes the change in production work from physical tasks to more activities that include planning and creative components and, at the same time, support these human cognitive abilities through the use of learning systems based on Industrie 4.0 technologies in production. Production work will continue to change significantly through collaboration with artificial intelligence within work processes and in many cases will mean dynamic handovers between system and human. [14–16]

2.2 Framework for Cognitive Production Work 4.0

Frameworks originate from information systems architecture and deal with the definition and control of logical structures for the integration of components within a system. [17,18] A framework "is a tool which can be used for developing a broad range of different architectures. It describes a method for designing an information system in terms of a set of building blocks, and for showing how the building blocks fit together. It contains a set of tools and provides a common vocabulary. It also includes a list of recommended standards and compliant products that can be used to implement the building blocks." [19] According to [17], frameworks can be characterized by the following features: model characteristics (representation, abbreviation, pragmatism), logical structure, methodical approach, intention/target group. [17]

The *Framework for Cognitive Production Work 4.0* is a development by researchers from the Future Work Lab and has been conceptualized with a domain-specific five-staged development process and according to eleven design principles in [7]. It is set up to serves six distinct aims, including both theoretical and practical elements. [16] The thematic framework elements are structured according to the widely-accepted *Human-*

Technology-Organization (HTO) classification. [20] Following this, the first version of thematic framework elements on two sub-levels and a more complex relation model has been developed in [7]. Figure 2 shows the elements of the first framework level divided into the HTO structure. See [7] and [16] for more details on the framework conceptualization and development as well as the expected benefits of the approach.

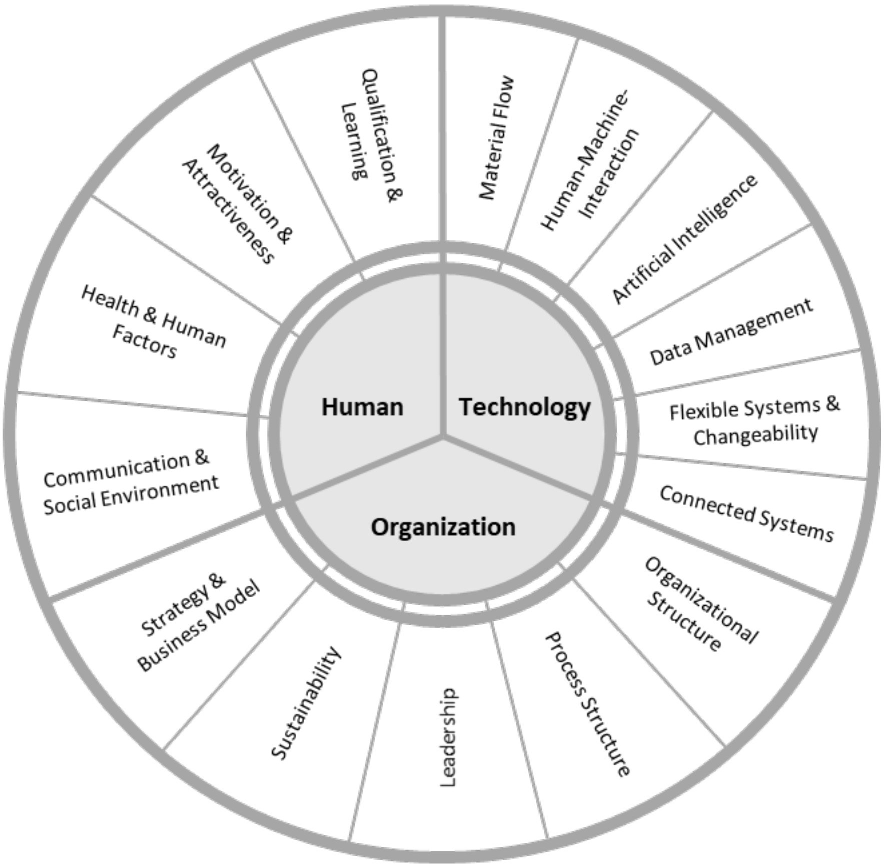


Figure 2: Framework for *Cognitive Production Work 4.0* [16]

The above presented structure of the framework is of mainly theoretical nature. In order to increase the applicability of the framework, the development team initiated multiple activities for practical examples of the application. Typical patterns of manufacturing value streams have been analyzed and transferred into schematic process structure for the direct application in production environments. As a visually supported action for the communication of the framework elements, graphical mood boards for three lead topics of *Cognitive Production Work 4.0* have been developed and communicated to users of the framework. [21] Another practical application of the framework can be observed in the demonstration world of the Future Work Lab. Here, a set of more than 40 use cases has been allocated to seven categories, which form a specific application layer of the framework. The classification approach is laid out in detail in [22]. The seven categories of the application layer are listed below:

- Ergonomics and safety
- Qualification and learning on the job
- Connected manufacturing systems
- Digital assistance
- Human-robot collaboration
- Intelligent machinery and systems
- Virtual engineering and planning [22]

3. Practical orientation guide for Industrie 4.0 best practices

The aim of this paper is to develop a suitable orientation guide for Industrie 4.0 best practices to support industrial companies in the process of their use case implementation. Therefore, the corresponding thematic structure (section 3.1) and process for the use of the orientation guide (section 3.2) are being described in this chapter. The limitations and further potentials are shortly discussed in section 3.3.

3.1 Thematic structure and categories

The thematic structure can be built upon the existing structure of the application layer of the *Framework for Cognitive Production Work 4.0* from 2.2. In order to address the needs of industrial companies, the existing seven categories needed revision for the intended purpose. The revision of the categories included literature analysis, data analysis of existing use case data bases and expert workshops. As a result, eight distinct categories have been identified for the practical orientation guide, which will be explained in subsections of this chapter.

For an intuitive understanding of the thematic fields, exemplary use cases were added to each category. The use cases were analyzed from different Industrie 4.0 data sets, including use cases from European level (*European Factories of the Future Research Association*) [23], German federal level (*Plattform Industrie 4.0*) [24], German regional level (*Allianz Industrie 4.0 Baden-Württemberg*) [25], and the German federal pilot project *Future Work Lab* [26]. Figure 3 shows the resulting composition of categories and exemplary use cases from the Future Work Lab. In the following, each category will be briefly defined and explained with a corresponding example.

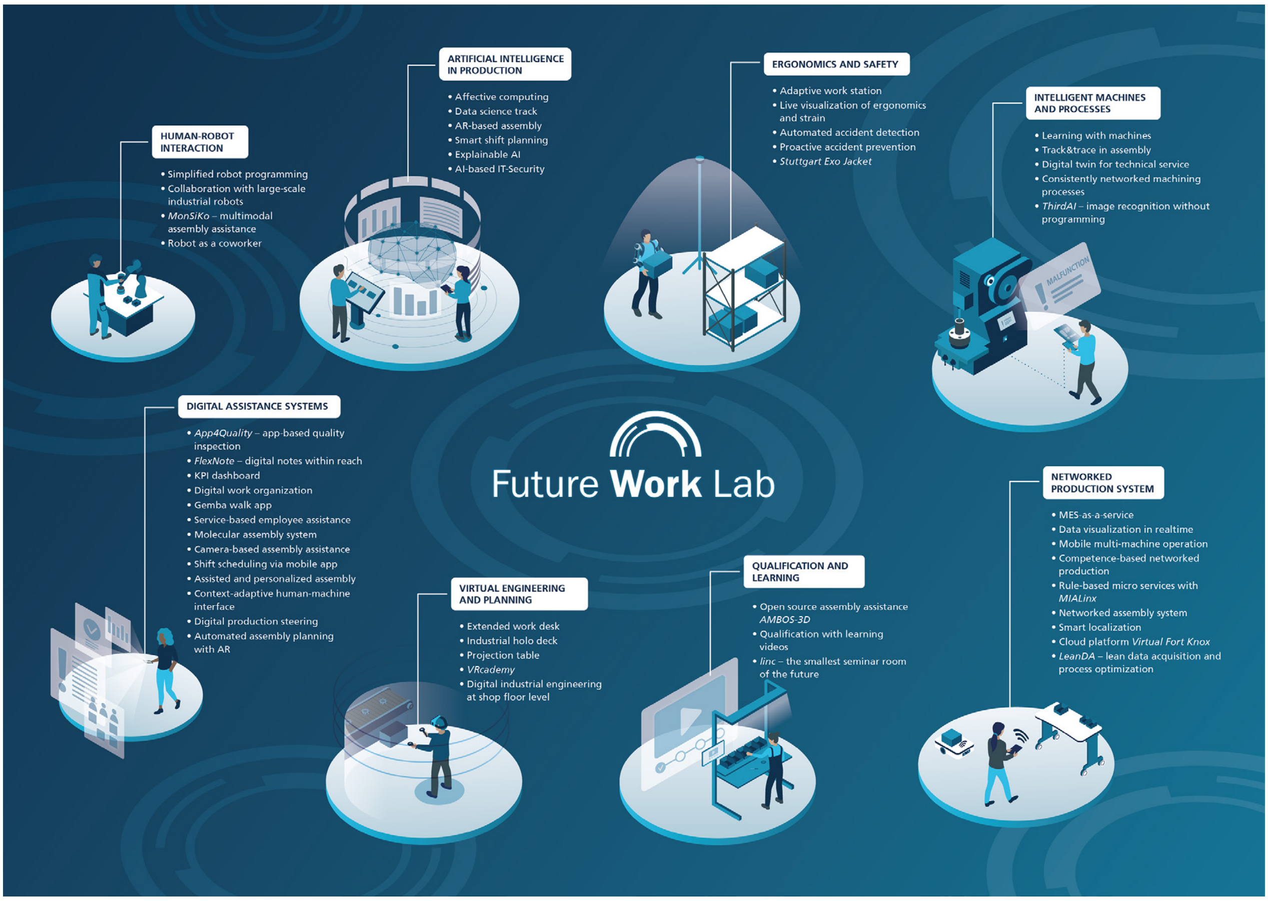


Figure 3: Practical orientation guide for Industrie 4.0 best practices with exemplary use cases from the Future Work Lab
© Valentina Klepikova, Fraunhofer IAO [26]

3.1.1 Human-robot interaction

Many current and future applications in the field of Industrie 4.0 will further reduce the distance (in terms of physical and work-related content) between humans and robots. In many cases, the intention is not to replace human work with technology, but to achieve meaningful cooperation and collaboration between workers and technology in order to enhance human capabilities. Robots are not solely placed in production cells behind fences, but are integrated into collaborative work activities. [27] E.g. sensitive lightweight robots and virtual safety fences enable direct cooperation between humans and robots.

3.1.2 Ergonomics and safety

The reduction of physical strain at work and the increase of occupational safety can be enhanced by intelligent sensor technology and the use of actuators. This category includes the organizational and technologically supported optimization of work task for the benefit of human health. Industrie 4.0 enables multiple innovative solutions based on cyber-physical systems, such as the use of passively or actively driven exo skeletons.

3.1.3 Intelligent machines and processes

Set-up and other activities during machine downtime often take up a lot of time without generating any added value. Intelligent machines can provide relevant data for workers and thus reduce downtime. Legacy machines can be retrofitted and thereby integrated into digital networks of machine parks on cloud platforms. Intelligent solutions combine the connectivity of machines and availability of data into process knowledge based on descriptive, diagnostic, predictive or prescriptive analytics. E.g. a digital twin model can be of such use.

3.1.4 Digital assistance systems

Digital assistance systems are commonly defined as computer-based systems for the support of humans with contextual information during three activities: the perception of information, decision making and the execution of tasks. [28] Humans can be assisted by a wide variety of technological applications. E.g. context information can be supplied via a mobile app, in form of a visualization on a screen or directly at the workplace.

3.1.5 Virtual engineering and planning

In addition to direct activities on the shop floor, engineering and planning are also changing significantly with Industrie 4.0. Virtual systems and tools such as Augmented Reality devices are opening up new possibilities in product development, production planning and many other areas. E.g. physical cardboard engineering can be fully transferred into virtual cardboard engineering with immersive features and underlying process models.

3.1.6 Qualification and learning

Industrie 4.0 technologies enable the integration of learning content directly at the workplace and in relation to individual needs of employees in production. Modular learning content can be provided via digital platforms and consumed independently from rigid structures or limited resources. New forms of qualification methods and tools are assigned to this category, e.g. for life-long learning, re- and upskilling of workers and learning near or on-the-job.

3.1.7 Networked production system

The networking of machines, workstations and tools in production is constantly evolving and offers both productivity and efficiency potential. Along with Industrie 4.0, concepts of networked and interoperable entities such as the *(Industrial) Internet of Things* and *cyber-physical production systems* enable new capabilities within factories. Networked entities also include the potential for the collection, use and analysis of large amounts of data.

3.1.8 Artificial intelligence in production

Artificial intelligence in production has been added as the eighth category in order to address the rapidly growing group of emerging solutions from machine learning algorithms and neural networks based on production data. Use cases vary from typical application fields such as maintenance, logistics, automation technology, quality management and control, product and process development, process planning, and resource planning. With regard to human emotions in production, an exemplary use case from the Future Work Lab is the implementation of affective computing in production, which means that workers' emotions are analyzed for the detection of flow phases and corresponding adaptions in task management. [26]

3.2 Application process

The use of the orientation guide is dependent on the needs of industrial users. For companies in the early orientation phase which have not yet started with the development of a use case, a top-down process is recommended in order to first identify a thematic category of applications and then seek for a suitable solution out of the suggested group of existing use cases. For companies in the use case development or implementation phase, a bottom-up process promises several advantages. Matching the own use case with existing best practices of comparable use cases of the orientation guide allows valuable insights. For example, this concerns technologies used in development and the adaptation of processes in implementation. Both processes are combined in Figure 4.

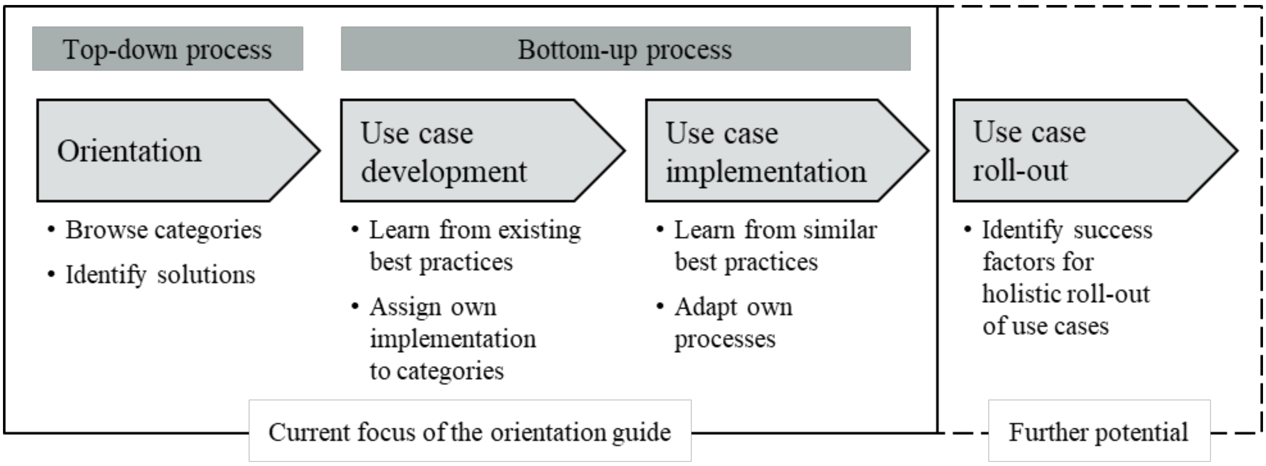


Figure 4: Process for the use of the practical orientation guide for Industrie 4.0 best practices in different phases

3.3 Limitations and potentials

In a first theoretical evaluation, the orientation guide has been found eligible for the thematic categorization of Industrie 4.0 use cases. When tested with new use cases in a practical evaluation, a small fraction of use cases might not be assignable to one of the predefined thematic fields. This limitation is calculated by design, but should not occur often. The orientation guide is expected to be even more valuable to users after further iterations including user feedback. Further potential of the orientation guide could be leveraged by the extension of the guide by integration of pre-developed value stream patterns as well as the expansion of the use case descriptions to include recommendations and success factors for the use case roll-out. However, this extension is planned in a future phase of development. The applicability of the orientation guide is still to be proved in a practical application case, which will be an international study described in the next chapter.

4. Outlook: International study on Industrie 4.0 best practices and success factors

The orientation guide supports its users to get a structured overview of possible Industrie 4.0 use cases and allows to assign own applications to the given thematic categories. The definition of the thematic categories is mainly based on best practices from German industry. In order to include global best practices, an international study on Industrie 4.0 best practices has been planned. This study will analyze best practices of manufacturing companies from seven different nations located in Europe, North America and Asia. Whereas the practical aim of the study is to provide companies from the region of Baden-Württemberg with insights into international best practices and success factors, the scientific aim of the study is also to validate the thematic categories of the orientation guide in an international context.

In order to achieve the aims of the study, a two-stage mixed-methods research design has been set up. The first stage is a standardized online questionnaire and the thematic categories of the orientation guide have been incorporated into the questionnaire. Table 1 gives an overview about the structure and content of the questionnaire. Based on the results of the questionnaire, the second stage of the study consists of expert interviews with representatives of the best practice companies involved. The expert interviews are conceptualized as semi-structured guideline interviews. The study is scheduled to be conducted from May until September 2021. The results will then be published.

Table 1: Structure and content of the questionnaire

Questionnaire section	Content	Exemplary items
Company demographics (3 items)	Brief characterization of companies for comparison	Industry sector; number of employees
General preconditions (16 items)	General preconditions with direct impact on the implementation of Industrie 4.0 use cases	Maturity level of Lean Production; existing infrastructure
Best practice – brief description (4 items)	Short summary of the best practice and assignment to the thematic categories of the orientation guide	Title of best practice; category according to the practical orientation guide
Best practice – situation before implementation (3 items)	Determination of the situation before the best practice was implemented	Know how in the field of Industrie 4.0; amount of implemented Industrie 4.0 use cases
Best practice – implementation phase (9 items)	Data from the implementation phase of the best practice project in order to derive success factors	Costs; external partners
Best practice – organization (5 items)	Analysis of the organizational characteristics of the best practice project	Involved employees/positions; size of the project team
Results (5 items)	Assessment of qualitative and quantitative outcomes of the best practice	Return on investment; quantitative and qualitative outcomes

5. Conclusion

In this paper a useful and practical orientation guide for Industrie 4.0 best practices is presented. The developments are based on a previously published and theoretically proven framework. The orientation guide provides any manufacturing company with suitable categories and a corresponding application process for the establishment of own Industrie 4.0 use cases. Furthermore, the assignment of over 40 use cases to eight thematic categories was successfully performed and proves as a first validation of the structure.

The application of the practical orientation guide for Industrie 4.0 best practices in an international study is already scheduled and prepared. This extensive validation action will generate further insights on the applicability of the structure and show potentials for iterations of the model at the same time.

Acknowledgements

This paper is part of two research projects. The practical orientation guide and international study are funded by the regional project “100 Orte für Industrie 4.0 in Baden-Württemberg”, while the underlying framework has been developed in the German Federal research project Future Work Lab.

The “100 Orte für Industrie 4.0 in Baden-Württemberg” project is funded by the Baden-Württemberg Ministry of Economy, Labour and Housing.

The Future Work Lab pilot project 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 implemented by the Project Management Agency Karlsruhe (PTKA).

References

[1] Bauer, W., Hämmerle, M., Bauernhansl, T., Zimmermann, T., 2019. Future Work Lab: The workplace of the future, in: Neugebauer, R. (Ed.), Digital Transformation. Springer Vieweg, Berlin, pp. 171–187.

[2] Bauernhansl, T., 2017. Die Vierte Industrielle Revolution: Der Weg in ein wertschaffendes Produktionsparadigma, in: Vogel-Heuser, B., Bauernhansl, T., Hompel, M. ten (Eds.), Handbuch Industrie 4.0 Bd.4. Springer, Berlin, Heidelberg, pp. 1–31.

[3] Neugebauer, R. (Ed.), 2019. Digital Transformation. Springer Vieweg, Berlin.

[4] Bauernhansl, T., Schatz, A., Jäger, J., 2014. Komplexität bewirtschaften – Industrie 4.0 und die Folgen. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 109 (5), 347–350.

[5] Dombrowski, U., Richter, T., Krenkel, P., 2017. Interdependencies of Industrie 4.0 & Lean Production Systems: A Use Cases Analysis. Procedia Manufacturing 11, 1061–1068.

[6] Schumacher, A., Nemeth, T., Sihn, W., 2019. Roadmapping towards industrial digitalization based on an Industry 4.0 maturity model for manufacturing enterprises. Procedia CIRP 79, 409–414.

[7] Schumacher, S., Pokorni, B., Himmelstoß, H., Bauernhansl, T., 2020. Conceptualization of a Framework for the Design of Production Systems and Industrial Workplaces. Procedia CIRP 91, 176–181.

[8] Plattform Industrie 4.0, 2021. Was ist Industrie 4.0? https://www.plattform-i40.de/PI40/Navigation/DE/Industrie40/WasIndustrie40/was-ist-industrie-40.html.

[9] Schumacher, S., Bildstein, A., Bauernhansl, T., 2020. The Impact of the Digital Transformation on Lean Production Systems. Procedia CIRP 93, 783–788.

[10] 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).

[11] Issa, A., Hatiboglu, B., Bildstein, A., Bauernhansl, T., 2018. Industrie 4.0 roadmap: Framework for digital transformation based on the concepts of capability maturity and alignment. Procedia CIRP 72, 973–978.

[12] DIN/DKE, 2018. Deutsche Normungsroadmap Industrie 4.0: Version 3, Berlin.

[13] Bloching, B., Leutiger, P., Oltmanns, T., Rossbach, C., Schlick, T., Remane, G., Quick, P., Shafranyuk, O., 2015. Die Digitale Transformation der Industrie. https://bdi.eu/media/presse/publikationen/information-und-telekommunikation/Digitale_Transformation.pdf.

[14] Bauernhansl, T., Tzempetonidou, M., Rossmeissl, T., Groß, E., Siegert, J., 2018. Requirements for designing a cyber-physical system for competence development. Procedia Manufacturing 23, 201–206.

[15] Rothe, I., Wischniewski, S., Tegtmeier, P., Tisch, A., 2019. Arbeiten in der digitalen Transformation – Chancen und Risiken für die menschengerechte Arbeitsgestaltung. Zeitschrift für Arbeitswissenschaft 73 (3), 246–251.

[16] Schumacher, S., Pokorni, B., 2020. Framework for Cognitive Production Work 4.0: Conceptualization and Model Approach for the Production System of the Future within the Future Work Lab. WT Werkstattstechnik 110 (3), 108–112.

[17] Matthes, D., 2011. Enterprise Architecture Frameworks Kompendium. Springer, Berlin, Heidelberg.

[18] Zachman, J.A., 1987. A framework for information systems architecture. IBM Systems Journal 26 (3), 276–292.

[19] Van Haren, 2007. The Open Group Architecture Framework TOGAF 2007 edition. Van Haren Publ, Zaltbommel.

[20] Ulich, E., 2013. Arbeitssysteme als Soziotechnische Systeme - eine Erinnerung. Psychology of Everyday Activity 6 (1), 4–12.

[21] Future Work Lab, 2021. Framework Cognitive Production Work 4.0. https://futureworklab.de/de/Framework.html.

[22] Findeisen, S., Körting, L., Schumacher, S., Eusterwiemann, T., Hämmerle, M., Pokorni, B., 2019. Classification Approach for Use Cases Within a Demonstration Factory Environment. Procedia Manufacturing 39, 106–116.

[23] European Factories of the Future Research Association, 2021. EFFRA Innovation Portal - Results, demos etc. overview. https://portal.effra.eu/results/list/result.

[24] Plattform Industrie 4.0, 2021. Industrie 4.0 Landkarte. https://www.plattform-i40.de/PI40/Navigation/DE/Angebote-Ergebnisse/Industrie-4-0-Landkarte/industrie-4-0-landkarte.html.

[25] Allianz Industrie 4.0 in Baden-Württemberg, 2021. 100 Orte für Industrie 4.0 in BW. https://www.i40-bw.de/de/100-orte/.

[26] Future Work Lab, 2021. Demonstrator World. https://futureworklab.de/en/demonstrator-environment.html.

[27] 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/.

[28] Link, M., Hamann, K., 2019. Einsatz digitaler Assistenzsysteme in der Produktion. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 114 (10), 683–687.

Biography

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.

Bastian Pokorni (*1985) is a researcher at Fraunhofer Institute for Industrial Engineering IAO in Stuttgart and head of the department for Connected Manufacturing Systems.

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.

Andreas Bildstein, Ph.D. (*1971) is head of 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.

Dr.-Ing. Moritz Hämmerle (*1983) is a researcher at the Fraunhofer Institute for Industrial Engineering IAO in Stuttgart and director of the department Cognitive Engineering and Production.

2nd Conference on Production Systems and Logistics

Tasks and Hurdles of Digital Transformation in Companies – a Literature-Review

Hanna Middendorf[1], Sandra Kaczmarek[2], Natalia Straub[1], Alexander Kreyenborg[1], Maximilian Austerjost[3], Michael Henke[2]

[1]*Chair of Enterprise Logistics, TU Dortmund University, Dortmund, Germany*

[2] *Chair of Enterprise Logistics, TU Dortmund University / Fraunhofer Institute of Material Flow and Logistics IML, Dortmund, Germany*

[3] *Insitute of Material Flow and Logistics IML, Dortmund, Germany*

Abstract

Digital Transformation in logistics and supply chain management is a challenge and a big chance for companies at the same time. It can open new business areas by providing new products and services. However, as promising as the digital transformation looks, the challenges for a successful transformation are not to underestimate. Many companies fail during the digital transformation process.

This paper aims to shed light on why some companies have difficulties with digital transformation and others not by looking at the hurdles and responsibilities involved. By comparing the limitations and obligations, this paper identifies gaps and ambiguities that might lead to failure.

A structured qualitative literature review presents the status quo of the actual scientific discourse in digital transformation, sums up tasks and barriers of the digital transformation in companies, and derives initial recommendations for action.

This paper shows that a digital transformation's hurdles and barriers are multifaceted and can be located on different enterprise levels. Enterprise culture and the top management's awareness and capabilities are essential for digital transformation. The same applies to the tasks of a digital transformation.

Keywords
Digital Transformation; Logistics; Supply Chain Management; Barriers; Tasks; Recommended Action

1. Digital Transformation in Logistics and Supply Chain Management

Digital transformation (DT) is a phenomenon that occurs equally in all industries. It offers companies the opportunity for new product and service innovations and the optimization of processes and workflows by using new means of digital technologies. This rapid technological progress affects all areas of a company, including logistics and supply chain management (LSCM). As a result, companies are under enormous pressure to change, which means that DT is no longer an optional process, but rather an obligation for management to address and implement to an appropriate degree. [1–3]

DT is becoming a strategic factor for competitive business purpose [4–7]. It is assumed that only those companies will survive in the new business environment that prepare for and accept the transformation process [8]. For LSCM, this means transforming value creation through digital technologies, changing and adapting strategies and techniques, and integrating enablers such as innovation and leadership to increase agility and realize higher productivity and customer-centric supply chains [9]. Due to the financial volume and the high number of employees, LSCM is one of the most important economic sectors in Germany [10,11], and is therefore considered in this paper. One challenge is the increasing complexity along supply chains, which requires a structured approach to support digital transformation management [12,13].

DT is a continuous and complex undertaking that offers opportunities to change a company and its operations significantly [4,6], and it is argued in [14] that very few conventional techniques will be sustainable in the future to support this radical change process. For this reason, it is also necessary to examine the critical capability of entrepreneurial thinking and action with regards to modern working methods and new technical possibilities.

According to Bucy et al., many companies will fail on this path because 70% of all DT are not successfully implemented [15]. The implementation process takes much longer and is more complex than expected, although technological inventions are made and corresponding roadmaps for implementation of innovations are planned [16–18].

By comparing and analyzing tasks and hurdles from scientific papers, it is possible to draw first conclusions as to why a successful DT is perceived as such a challenge. This paper then gives practical conclusions for LSCM. Specifically, the question was explored as to why the majority of companies (70%, see above) is having difficulties implementing DT while a significant share (30%) is successful.

2. Tasks and Barriers of Digital Transformation

To achieve a comprehensive understanding of the complex phenomenon of DT in the field of LSCM, this paper uses a structured qualitative literature review according to Xiao and Watson [19]. Table 1 shows the information on how the literature review was conducted.

Table 1: Overview of literature review

Platforms used		Google Scholar, EBSCOhost, database of the University library			
Search string		"digital transformation" OR "digitization" OR "digitalization" AND "barriers" OR "hurdles" OR "obstacles" OR "challenges" AND/OR "task" OR "mange" OR "navigate"			
Publication period		2000 till now			
Initial paper volume		1.210			
Selection criteria		Ranking (A+ to D); impact factor, German or English language; relevance for the topic			
Papers on Digital Transformation		Models	Studies	Framework	Literature Reviews
A	[20,4,21–25]		6		1
B	[26–38]		6	6	1
C	[39–43]	2		5	
D	[44–49]	2	1	3	
Impact Factor	[50–59]		4	5	1
Without Ranking	[60–70]	3	3	5	
Overall	54	7	20	23	3

The literature review served to approach the topic of managing DT and to identify why some companies successfully implement this change and others do not. The findings, mainly from management research, are focused on the field of LSCM in the discussion in order to systematically prevent misleading interpretation due to specific challenges in other sectors. Based on the extensive literature review, eight categories were identified that reflect the digital change in companies comprehensively: Corporate culture, business model, strategy, organisation, employees, technology and data, resources (monetary) and top management. These categories were summarized with a focus on brevity.

Corporate Culture: An important task is the adaptation of the corporate culture [39,4,27,28,61,46,47,35,36,48,70] to meet the requirements and implementation of the DT. At the same time, however, this is also a major hurdle, because many companies find it challenging to implement the change successfully [54,63,64,31,47,37,68,69]. Besides, a risk-averse corporate culture is an obstacle to successfully implementing the DT [28,57].

Business Model: The business model plays an essential role in the tasks of a DT, because it (often) needs to be adapted and changed to address new tasks and processes of the DT. It is recommended to make the business model more agile and leaner. [44,39,26,40,30,32,43,47,34,70] The barriers that need to be overcome in the business model are described by changing customer needs and expectations and outdated business models [44,39,45].

Strategy: In this category, several tasks need to be addressed for a successful management of the DT: Companies have to formulate and implement a resilient digital/DT strategy – at the best, the digital strategy is embedded in the corporate strategy [39,26,40,28,32,65,42,66,47,35,37,24,25,70]. A clear and understandable and transparent DT vision is defined [27,40,28,61,58,33]. Governance should be established for constant reflection and evaluation [30,47]. The role of a Chief Digital Officer (CDO) or Digital Leader should be introduced or at least considered as well as a DT team to strengthen the digital strategy [40,21,30,62,31,47,37,24,48,25]. One hurdle listed in this context is that it is a hindrance if there is no strategy and no understanding of a DT [54,28,57,45,68].

Organisation: Within an organization's framework, the company has to be leaner and more flexible to meet the DT requirements [51,39,4,52,56,40,31,36]. One task is to actively promote cooperation within the organization, especially with cross-departmental and cross-team work. The silo mentality of departments must be reduced or eliminated. [39,26–28,61,31] The orientation of the organization, as well as processes, interfaces and role and job profiles of the employees must be adapted within the company so that a digital ecosystem can emerge. [20,44,30,65,67,34,48]. The following barriers could be identified in this category: The DT challenges the organization's fundamental and grown structure [20,48]. On the other hand, the existing organisational structure may hinder change. If a company is digitally inexperienced, a DT can be a major challenge by itself [55,45,63,23]. In general, the organization must be willing to embrace change [39,54,57,65,37]. Strict separation of teams and tasks is a hindrance as interdisciplinary work is identified to necessary [53,41,59].

Employees: The tasks related to employees for the management of a DT are identified in the literature as follows: Employees must be involved in the transformation process from the start to, e.g., convert their curiosity into new ideas or suggestions [4,56,40,43,33,48,70]. Companies should promote, develop and gain digital skills and optionally implement new professional roles [52,26,40,28,30,61,58,24]. Furthermore, workplaces, working hours and work content changes (lifelong learning, data processing, remote work) and training must be adapted accordingly [50,39,52,60,31,46,65,43,47]. Barriers are reluctance to changes of the staff [27,59], missing digital skills, know-how and expertise [54,28,57,63–65,59,68], and fundamental new requirements for the qualification of employees [52,55].

Technology and Data: In technology and data, the following tasks for a DT have been identified: The development of an IT infrastructure, e.g. by setting up a platform and/or database [26,40,31,65,35,70]. The management and asset-oriented use of data [26,40,31]. Big data analysis, data mining and data analytics [40,28]. A suitable technology selection, management, implementation, and process integration [30,31,70]. The main hurdles are as follows: Data security [53,57,63,59,47], uncertainty what technology to choose [53,45,66], and an insufficient IT structure [71,63,64].

Resources (monetary): In the field of resources there are some tasks to address; Budget for advancing a DT (testing new ideas, new forms of software etc.) should be reserved [4,28] and investments must be designed to suit the company [31,32,70]. Therefore, the main obstacle is the lack of insufficient financial resources [71,57,63,68,69], lack of resources [53,59], and also when companies misconceive a DT for a one-time investments and thus do not budget corresponding resources [55,42,35].

Top management: In the category of top management, the following tasks for a DT were identified: Top management should define and communicate a digital strategy and also name the most important activities [41,42,48,49,70]. The top management should regularly reflect on the necessity of their own mindset and thinking [50,52,29,22,67,34,38,25]. The top management team has to involve internal and external support and consultancy for a proper DT [34,24,25]. The management should encourage and support employees and act as a role model (live and support digital vision) [29,61,23,33,67,38]. Entrepreneurial action (prioritizing tasks, allocating resources, weighing up opportunities and risks) should be actively taken up and pursued [27,29,23,34,25]. Two major aspects were mentioned in the literature as obstacles to DT: Feeling powerless, overwhelmed with deciding what to do – Leadership for DT differs from conventional change [71,29,57,30,63,23,48]. If leadership does not drive DT, it will fail [71,63,68].

3. Why do some companies face hurdles

The tasks and barriers of a DT management are multidimensional and affect the technical decisions and the entire organization's change, including the corporate culture [39,4,27,28,61,46,47,35,36,48,70]. Companies embarking on the path of DT are leaving familiar tracks behind and have to reinvent themselves [44,40,30,32,34].

According to Schoemaker et al. ordinary skills used in established companies are not suitable for the implementation of a DT [34]. Since the DT has a disruptive character, it cannot be adequately responded to with conventional skills. Therefore, there is a need to approach the phenomenon of DT through dynamic capabilities [72].

Based on the MTO concept according to Ulich [73], a design approach is characterised here that encompasses the characteristics of "role of man", "technology" and "organization" (the so called MTO model). DT increases the quality and quantity of real-time data available to companies and offers them new opportunities. Using analytical methods and artificial intelligence, fundamental information about one's own business processes can be provided, e.g., process, environmental and status data, which is a valuable asset. This can lead to a self-optimised organisation that enables autonomous and timely adaptation to changing conditions in the business environment. [74] For this reason, the MTO model is extended by the feature "information". Based on the results of the literature review and the MTOI approach, critical factors that lead to hurdles of a DT can be discussed.

For the successful implementation of the DT, it is important to set targets for a realisation in the application domain of supply chain management, among others suitable production and logistics systems. If no or unsuitable targets are set, this can hinder the exploitability and development of a successful new business model. It is necessary to design new business models with the permanent business processes and value creation processes. Their design should enable the company to exploit technological means economically [39,40,30]. However, this can only be achieved if the customers' new requirements and needs are correctly recognized and appropriately interpreted [39,66].

After defining the targets companies can enter the planning. This is where the characteristics of the MTOI approach become relevant for detailed consideration, because all four features have a significant influence on the processes of designing and planning production and logistics systems. *Technology:* In this phase, decisions regarding the technology should be made. Logistics and supply chain companies need to adapt their IT infrastructure and make a choice of appropriate technologies to go into the design and planning of production and logistics systems. This is already the first major challenge, as digitally inexperienced companies have difficulties making the right choice of technologies to strengthen their business model [31,70]. *Information:* When considering the IT infrastructure, initial thoughts should be given to how the (newly available) data should be used and exploited. *Organisation:* To successfully realize and implement a DT, the organisation's fundamental structures must change. Just because the previous structure has proven to work in the past, a change of the organization towards an agile, flexible and lean structure [4,31,36] is now necessary to identify changes early on and to be able to react to them flexibly and appropriately [34] to design and plan production and logistics systems. Likewise, the division of work by (organizational) departments should be dissolved, and interdisciplinary work should be promoted instead [39,27] in order to be able to react more quickly to changes. All this requires a change in corporate culture. The impact of the DT on corporate culture is just as radical as in all other areas. The corporate culture must change in order to promote flexibility and innovation [46,48]. *People:* Radical change can only be implemented successfully if employees are also willing to change. Therefore, employees should be involved in the transformation process and actively participate with curiosity and ideas [4,56,40,43,33,48,70]. The rejection and mistrust of employees [27,59] can represent an insurmountable hurdle for the company concerning the successful implementation of the DT. Employees must be willing to build digital skills and expertise and embrace the journey of change out of their conviction [52,56,61]. The role of employees is changing, as they are required to work in a more self-determined way and also to network independently with other employees and departments [52,56].

Planning is followed by the decision-making phase. This is about deciding whether and how the plans should be realised. To do this, the future technology portfolio, the role of people, organisational models, and the integration of information for the company must be developed. The initial use cases should be modelled explicitly, because companies that do not formulate detailed business cases are known to have difficulties describing and evaluating the exact economic benefits. [75]

The next step is the realization - the integration of new technologies into existing business processes. At this point critical factors like lack of data security [53,57,63,59,47] and an insufficient IT structure [71,63,64] can occur. Interfaces to the customer must be defined and technologies from different manufacturers need to be brought together. The organisational change needs to be driven actively and the radical change of a DT must be addressed appropriately. This also means well-structured human resource management to drive personal development and employee competencies along with organizational change and thus implement necessary changes in the work system.

Monitoring DT is necessary to measure its success and to keep an eye on the strategic fit of a company's entire structure. Companies will have difficulties if they assume that the DT can run alongside day-to-day business like conventional projects. Monitoring is an integral part of company-wide strategic early detection. The aim of this early detection is to provide timely information about changes in the entire field of the company in order to identify possible opportunities and risks at an early stage. Governance should be established for constant reflection and evaluation [30,47].

To engage in the process of DT, there are several factors that accompany the entire process and contribute significantly to its success or failure. It is necessary to have a well-developed, resilient and clearly formulated digital strategy, which in the best case is integrated into the overall corporate strategy [39,40,28,32,65,42,66,47,35,37,24,25,70]. Companies make the biggest mistake by viewing the DT as a project and therefore lack an understanding of the DT and accordingly no strategy is formed [28]. Another factor is the allocation of resources, which can be challenging for companies. This is because the distribution of financial resources does not follow the standard project approach. Funds must be made available to test, try out and implement new ideas without a defined business case in the beginning. Furthermore, companies need to be aware that the decision for a DT means a long-term investment [55,42,35]. A third factor is the top management. The leadership of a DT is different from that of a normal change [76]. Due to all the changes and new requirements of a DT, the role of leadership also changes. They need to become drivers of the DT [71,63,68]. A particularly great challenge for leadership is to reflect on themselves constantly and to change their own mind-set and thinking [50,52,29,22,67,34,38,25].

4. Theoretical and managerial implications and further research

The reasons why companies have difficulties with DT are many and varied and are manifold and based primarily on the disruptive nature of the change. In short, the extent of the DT is simply underestimated by the more struggling companies. Based on the literature review, it can be tentatively (!) concluded that companies with a successful DT reshape the purpose and the setup of their venture. This hypothesis has to be confirmed, and findings of management research in the field of LSCM should be investigated in depth. One approach that should be considered further is the concept of dynamic capabilities. With the help of this theoretical approach, further recommendations for action can be derived, explained and defined. Further research should look more closely at roles and hierarchies of companies in LSCM because a radical change process means that traditional job profiles will also change. In LSCM, the extent to which roles change or are replaced by other roles must be examined at this point. It also should be investigated how top management acts more entrepreneurially because this could be a key capability in DT and organizational change. It makes sense to look at how established and structurally set companies can realize change and which influencing factors play a key role. Long-term studies that accompany the process of change in LSCM would be useful to obtain valid results.

Acknowledgements

This research and development project was funded by the German Federal Ministry of Education and Research (BMBF) within the “Zukunft der Arbeit” Program (02L18A190) and implemented by the Project Management Agency Karlsruhe (PTKA). The authors are responsible for the content of this publication.

References

[1] Grivas, S.G., Graf, M., 2020. Digitale Transformation – Transformation der Unternehmen im digitalen Zeitalter, in: Gatziu Grivas, S. (Ed.), Digital Business Development, vol. 31. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 143–166.

[2] Moufaddal, M., Benghabrit, A., Bouhaddou, I., 2019 - 2019. Industry 4.0: A roadmap to digital Supply Chains, in: 2019 1st International Conference on Smart Systems and Data Science (ICSSD). 2019 1st International Conference on Smart Systems and Data Science (ICSSD), Rabat, Morocco. 03.10.2019 - 04.10.2019. IEEE, pp. 1–9.

[3] Preindl, R., Nikolopoulos, K., Litsiou, K., 2020. Transformation strategies for the supply chain: the impact of industry 4.0 and digital transformation. Supply Chain Forum: An International Journal 21 (1), 26–34.

[4] Berghaus, S., Back, A., 2017. Disentangling the fuzzy front end of digital transformation: Activities and approaches. Thirty Eighth International Conference on Information Systems.

[5] Gassmann, O., Sutter, P., 2016. Digitale Transformation im Unternehmen gestalten: Geschäftsmodelle, Erfolgsfaktoren, Handlungsanweisungen, Fallstudien. Hanser, München, 282 pp.

[6] Singh, A., Hess, T., 2017. How Chief Digital Officers promote the digital transformation of their companies. MIS Quarterly Executive 16 (1).

[7] Vetter, H.M., 2017. Dynamic Capabilities for the Digital Transformation: Scale Development and Empirical Investigation of Performance Implications. Apprimus.

[8] Marmolejo-Saucedo, J., Hartmann, S., 2020. Trends in digitization of the supply chain: A brief literature review. EAI Endorsed Transactions on Energy Web 7 (29), 1–7.

[9] Junge, A.L., Straube, F., 2020. Sustainable supply chains – digital transformation technologies' impact on the social and environmental dimension. Procedia Manufacturing 43, 736–742.

[10] Kübler, A., Distel, S., Veres-Homm, U., 2015. Logistikbeschäftigung in Deutschland: Vermessung, Bedeutung und Struktur. Fraunhofer Verlag, Stuttgart, 97 pp.

[11] Schwemmer, M. (Ed.), 2016. Die Top 100 der Logistik: Marktgrössen, Marktsegmente und Marktführer in der Logistikdienstleistungswirtschaft, Ausg. 2016/2017 ed. DVV Media Group Dt. Verkehrs-Verl., Hamburg, 400 pp.

[12] Clauss, A.-M., Schumann, C.-A., 2020 - 2020. A Systems Theory and Action Design Research Perspective on Supply Chain Collaboration in the Context of SCM 4.0, in: 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Cardiff, United Kingdom. 15.06.2020 - 17.06.2020. IEEE, pp. 1–8.

[13] Hompel, M. ten, Henke, M., 2017. Logistik 4.0 – Ein Ausblick auf die Planung und das Management der zukünftigen Logistik vor dem Hintergrund der vierten industriellen Revolution, in: Vogel-Heuser, B., Bauernhansl, T., Hompel, M. ten (Eds.), Handbuch Industrie 4.0 Bd.4. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 249–259.

[14] Mohr, T., 2020. Der Digital Navigator: Ein Modell für die digitale Transformation. Springer.

[15] Bucy, M., Finlayson, A., Kelly, G., Moye, C., 2016. The 'how'of transformation.

[16] Balakrishnan, R., Das, S., 2020. How do firms reorganize to implement digital transformation? Strategic Change 29 (5), 531–541.

[17] Glaser, B., Strauss, A.L., 1967. The Discovery of Grounded Theory. Aldine, Chicago.

[18] Zinder, E., Yunatova, I. Synergy for digital transformation: person's multiple roles and subject domains integration, in: , International Conference on Digital Transformation 2016, vol. 2016, pp. 155–168.

[19] Xiao, Y., Watson, M., 2019. Guidance on Conducting a Systematic Literature Review. Journal of Planning Education and Research 39 (1), 93–112.

[20] Baptista, J., Stein, M.-K., Klein, S., Watson-Manheim, M.B., Lee, J., 2020. Digital work and organisational transformation: Emergent Digital/Human work configurations in modern organisations. The Journal of Strategic Information Systems 29 (2).

[21] Haffke, I., Kalgovas, B.J., Benlian, A., 2016. The Role of the CIO and the CDO in an Organization's Digital Transformation.

[22] Li, L., Su, F., Zhang, W., Mao, J.-Y., 2018. Digital transformation by SME entrepreneurs: A capability perspective. Info Systems J 28 (6), 1129–1157.

[23] Ohain, B.P. (Ed.), 2019. Leader Attributes for Successful Digital Transformation.

[24] Soh, C., Yeow, A., Goh, Q., Hansen, R., 2019. Digital Transformation: Of Paradoxical Tensions and Managerial Responses.

[25] Yeow, A., Soh, C., Hansen, R., 2018. Aligning with new digital strategy: A dynamic capabilities approach. The Journal of Strategic Information Systems 27 (1), 43–58.

[26] Correani, A., Massis, A. de, Frattini, F., Petruzzelli, A.M., Natalicchio, A., 2020. Implementing a Digital Strategy: Learning from the Experience of Three Digital Transformation Projects. California Management Review 62 (4), 37–56.

[27] Eden, R., Jones, A.B., Casey, V., Draheim, M., 2019. Digital transformation requires workforce transformation. MIS Quarterly Executive 18 (1).

[28] Gurbaxani, V., Dunkle, D., 2019. Gearing Up For Successful Digital Transformation. MISQE 18 (3), 209–220.

[29] Hansen, A.M., Kraemmergaard, P., Mathiassen, L., 2011. Rapid Adaptation in Digital Transformation: A Participatory Process for Engaging IS and Business Leaders. MIS Quarterly Executive 10 (4), 175–185.

[30] Heavin, C., Power, D.J., 2018. Challenges for digital transformation – towards a conceptual decision support guide for managers. Journal of Decision Systems 27 (1), 38–45.

[31] Lindner, D., Ludwig, T., Amberg, M., 2018. Arbeit 4.0 – Konzepte für eine neue Arbeitsgestaltung in KMU. HMD 55, 1065–1085.

[32] Matt, C., Hess, T., Benlian, A., 2015. Digital Transformation Strategies. Bus Inf Syst Eng 57 (5), 339–343.

[33] Reijnen, C., Overbeek, S., Wijers, G.M., Sprokholt, A., Haijenga, F., Brinkkemper, S., 2018. A Shared Vision for Digital Transformation: Codification of The Operating Model Canvas Approach.

[34] Schoemaker, P.J.H., Heaton, S., Teece, D., 2018. Innovation, Dynamic Capabilities, and Leadership. California Management Review 61 (1), 15–42.

[35] Sebastian, I., Ross, J., Beath, C., Mocker, M., Moloney, K., Fonstad, N., 2017. How big old companies navigate digital transformation. MIS Quarterly Executive 16 (3), 197–213.

[36] Shaughnessy, H., 2018. Creating digital transformation: strategies and steps. SL 46 (2), 19–25.

[37] Singh, A., Klarner, P., Hess, T., 2020. How do chief digital officers pursue digital transformation activities? The role of organization design parameters. Long Range Planning 53 (3).

[38] Solberg, E., Traavik, L.E.M., Wong, S.I., 2020. Digital Mindsets: Recognizing and Leveraging Individual Beliefs for Digital Transformation. California Management Review 62 (4), 105–124.

[39] Berghaus, S., Back, A., 2016. Gestaltungsbereiche der Digitalen Transformation von Unternehmen: Entwicklung eines Reifegradmodells. Die Unternehmung 70 (2), 98–123.

[40] Gimpel, H., Hosseini, S., Huber, R.X.R., Probst, L., Röglinger, M., Faisst, U., 2018. Structuring Digital Transformation: A Framework of Action Fields and its Application at ZEISS. J. Inf. Technol. Theory Appl. 19 (1), 31–54.

[41] Horlacher, A., Hess, T., 2016 - 2016. What Does a Chief Digital Officer Do? Managerial Tasks and Roles of a New C-Level Position in the Context of Digital Transformation, in: 2016 49th Hawaii International Conference on System Sciences (HICSS). 2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI, USA. 05.01.2016 - 08.01.2016. IEEE, pp. 5126–5135.

[42] North, K., Aramburu, N., Lorenzo, O.J., 2020. Promoting digitally enabled growth in SMEs: a framework proposal. JEIM 33 (1), 238–262.

[43] Poethke, U., Klasmeier, K.N., Diebig, M., Hartmann, N., Rowold, J., 2019. Entwicklung eines Fragebogens zur Erfassung zentraler Merkmale der Arbeit 4.0. Zeitschrift für Arbeits- und Organisationspsychologie A&O 63 (3), 129–151.

[44] Baumöl, U., Bockshecker, A., 2018. Steuerung im Zeitalter der Digitalisierung mit dem Digital Business Management-Modell. CON 30 (5), 4–11.

[45] Hess, T., Barthel, P., 2017. Wieviel digitale Transformation steckt im Informationsmanagement? Zum Zusammenspiel eines etablierten und eines neuen Managementkonzepts. HMD 54 (3), 313–323.

[46] Mayer, J.H., Eßwein, M., Hornung, K., Steigner, M., 2019. The MENSCH in the Digital Age - Are next-generation managers triggering basic changes? CON 31 (5), 57–63.

[47] Reinhard, H., Rentz, R., Sommerfeld, T., 2020. Digitalisierung steuerbar machen. Control Manag Rev 64 (2), 46–51.

[48] Stöger, R., 2017. Umsetzung der Digitalisierung. zfo–zeitschrift für organisationsentwicklung 1.

[49] Walchshofer, M., Riedl, R., 2017. Der Chief Digital Officer (CDO): Eine empirische Untersuchung. HMD 54 (3), 324–337.

[50] Ameln, F. von, Wimmer, R., 2016. Neue Arbeitswelt, Führung und organisationaler Wandel. Gr Interakt Org 47 (1), 11–21.

[51] Bendel, A., Latniak, E., 2020. Soziotechnisch – agil – lean: Konzepte und Vorgehensweisen für Arbeits- und Organisationsgestaltung in Digitalisierungsprozessen. Gr Interakt Org 51 (3), 285–297.

[52] Boes, A., 2017. Qualifizieren für das Arbeiten im globalen Informationsraum. WSI 70 (2), 155–157.

[53] Cichosz, M., Wallenburg, C.M., Knemeyer, A.M., 2020. Digital transformation at logistics service providers: barriers, success factors and leading practices. IJLM 31 (2), 209–238.

[54] Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M., 2013. Embracing Digital Technology: A New Strategic Imperative. MIT Sloan Management Review., Massachusetts.

[55] Frank, M., Koldewey, C., Rabe, M., Dumitrescu, R., Gausemeier, J., Kühn, A., 2018. Smart Services – Konzept einer neuen Marktleistung. ZWF 113 (5), 306–311.

[56] Gergs, H.-J., 2019. Agilität und Organisationsentwicklung – Ziemlich beste Freunde? Gr Interakt Org 50 (2), 101–110.

[57] Harvad Business Review (Hrsg.), 2016. What is holding back the digial Revolution? Copyright © 2016 Harvard Business School Publishing.A Harvad Business Review Analytic Services Report, 1–8.

[58] Ivanov, D., Tang, C.S., Dolgui, A., Battini, D., Das, A., 2020. Researchers' perspectives on Industry 4.0: multi-disciplinary analysis and opportunities for operations management. International Journal of Production Research 53 (8), 1–24.

[59] Peillon, S., Dubruc, N., 2019. Barriers to digital servitization in French manufacturing SMEs. Procedia CIRP 83 (1), 146–150.

[60] Gebhardt, J., Grimm, A., Neugebauer, L.M., 2015. Entwicklungen 4.0–Ausblicke auf zukünftige Anforderungen an und Auswirkungen auf Arbeit und Ausbildung. Journal of Technical Education (JOTED) 3 (2), 45–61.

[61] Hoe, S.L., 2019. Digitalization in practice: the fifth discipline advantage. TLO 27 (1), 54–64.

[62] Issa, A., Hatiboglu, B., Bildstein, A., Bauernhansl, T., 2018. Industrie 4.0 roadmap: Framework for digital transformation based on the concepts of capability maturity and alignment. Procedia CIRP 72, 973–978.

[63] Lammers, T., Tomidei, L., Trianni, A., 2019 - 2019. Towards a Novel Framework of Barriers and Drivers for Digital Transformation in Industrial Supply Chains, in: 2019 Portland International Conference on Management of Engineering and Technology (PICMET). 2019 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA. 25.08.2019 - 29.08.2019. IEEE, pp. 1–6.

[64] Leipzig, T. von, Gamp, M., Manz, D., Schöttle, K., Ohlhausen, P., Oosthuizen, G., Palm, D., Leipzig, K. von, 2017. Initialising Customer-orientated Digital Transformation in Enterprises. Procedia Manufacturing 8, 517–524.

[65] Mielli, F., Bulanda, N., 2019 - 2019. Digital Transformation: Why Projects Fail, Potential Best Practices and Successful Initiatives, in: 2019 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA). 2019 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA), St Louis, MO, USA. 28.04.2019 - 02.05.2019. IEEE, pp. 1–6.

[66] Rautenbach, W.J., Kock, I.d., Jooste, J.L., 2019 - 2019. The development of a conceptual model for enabling a value-adding digital transformation: A conceptual model that aids organisations in the digital transformation process, in: 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Valbonne Sophia-Antipolis, France. 17.06.2019 - 19.06.2019. IEEE, pp. 1–10.

[67] Rotzinger, J., 2017. Unternehmen erfolgreich in (die) Zukunft führen. Gr Interakt Org 48 (4), 273–278.

[68] Solis, B., Li, C., Szymanski, J., 2017. The 2017 state of digital transformation. Altimeter Group 1 (1), 1–33.

[69] Wolf, M., Semm, A., Erfurth, C., 2018. Digital Transformation in Companies – Challenges and Success Factors, in: Hodoň, M., Eichler, G., Erfurth, C., Fahrnberger, G. (Eds.), Innovations for Community Services, vol. 863. Springer International Publishing, Cham, pp. 178–193.

[70] Yucel, S., 2018 - 2018. Modeling Digital Transformation Strategy, in: 2018 International Conference on Computational Science and Computational Intelligence (CSCI). 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA. 12.12.2018 - 14.12.2018. IEEE, pp. 221–226.

[71] Fitzgerald, M., Kruschwitz, N., Bonnet, D., Welch, M., 2014. Embracing digital technology: A new strategic imperative. MIT sloan management review 55 (2), 1.

[72] Warner, K.S.R., Wäger, M., 2019. Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. Long Range Planning 52 (3), 326–349.

[73] Ulich, E., 2020. Arbeitspsychologie. vdf Hochschulverlag AG.

[74] Schuh, G., Anderl, R., Gausemeier, J., Hompel, M. ten, Wahlster, W., 2017. Industrie 4.0 maturity index. Managing the digital transformation of companies. Munich: Herbert Utz.

[75] Henke, M., Besenfelder, C., Kaczmarek, S., Hetterscheid, E., Schlüter, F. Dortmund management model—a contribution to digitalization in logistics and supply chain management, in: Proceedings of the 9th International Scientific Symposium on Logistics: ISSL, pp. 13–14.

[76] Molina, K.-M., Kaiser, S., Widuckel, W., 2018. Kompetenzen der Zukunft-Arbeit 2030: als lernende Organisation wettbewerbsfähig bleiben. Haufe-Lexware.

Biography

Hanna Middendorf, M.A. (*1988) studied social science innovation research at the TU Dortmund and has worked as a research assistant at the Chair of Enterprise Logistics (LFO) since 2020.

Dipl.-Päd. Sandra Kaczmarek (*1986) studied educational science with a focus on vocational education at the TU Dortmund University and has worked as a research assistant at the Chair of Enterprise Logistics (LFO) and the Fraunhofer Institute for Material Flow and Logistics (IML) since 2014.

Dipl.-Log. Natalia Straub (*1978) studied logistics at the TU Dortmund and has worked as a Chief Engineer at the Chair of Enterprise Logistics (LFO) since 2011.

Alexander Kreyenborg, M. Sc. (*1988) studied automotive mechatronics at Munich University of Applied Sciences and industrial engineering at Münster University of Applied Sciences. Since 2016, he is research assistant at the TU Dortmund University at the Chair of Enterprise Logistics.

Maximilian Austerjost, M. Sc. (*1986) studied mechanical engineering with economics at the South Westphalia University of Applied Sciences in Meschede and the University of Paderborn. He has worked as a research assistant at the Chair of Enterprise Logistics (LFO) and the Fraunhofer Institute for Material Flow and Logistics (IML) since 2014.

Prof. Dr. Michael Henke (*1971) has been in the leadership of the Chair of Enterprise Logistics (LFO) and institute director of the Fraunhofer Institute of Material Flow and Logistics (IML) at the TU Dortmund University since 2013. He is also an adjunct professor for Supply Chain Management at the School of Business and Management at the Lappeenranta University of Technology in Finland since 2015.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Experimental Validation Of A Solidification Model For Automated Disassembly

Richard Blümel[1], Annika Raatz[1]
[1] Leibniz University Hannover, Institute of Assembly Technology, An der Universität 2, 30823 Garbsen, Germany

Abstract

Disassembly is a crucial step towards sustainable life cycle engineering. During the operation, assembly connections solidify to an unknown state, e.g. due to thermal or mechanical stress on the product. Therefore, disassembly forces are hard to predict. With regard to automated disassembly, this complicates the proper planning of disassembly times and tools. The uncertainties can lead to damage or destruction of the product, impeding regeneration. To tackle these problems, in earlier work, we proposed a solidification model, which enables planners to predict disassembly forces based on the products geometric properties and operational history without investigating the complex physical influences caused by the usage of the product. The disassembly of high-value capital goods like aircraft engines, in particular blade-disk connections, serves as an application case. Still, we were not yet able to validate the solidification model due to the lack of experimental reproducibility. In this work, we adapt the existing model of a solidified assembly connection created in prior work with an additional clamping force. The additional force aims to represent the solidification force. This can significantly increase reproducibility and reduce disturbances.

Keywords

Disassembly; Regeneration; Turbine Blades; Design of Experiments; Regression Model

1. Introduction

Disassembly is a process that is strongly characterized by uncertainties [1]. In contrast to assembly, essential process variables to plan an automated process such as the force to be applied and the duration of the process can only be estimated based on experience and cannot be determined precisely [2]. During operation, products are subjected to physical and thermal stress or corrosion. Thus properties, such as joining tolerances, used to plan assembly lose their validity. This way, detachable connections solidify into hardly or even non-detachable connections, resulting in unknown disassembly planning parameters. The Collaborative Research Centre (CRC) 871, "Regeneration of Complex Capital Goods", aims to develop a scientific basis for maintenance, repair and overhaul (MRO) using aircraft engines high-pressure turbine (HPT) blades as the focus of attention. The expenses for MRO of aircraft engines cause approximately one-third of the engine's operation costs [3]. New turbine blades can cost as much as $8,000 each [4], making regeneration appropriate. The main goal of the research within the CRC 871 is to restore or even improve the functional properties to save and regenerate as many worn components as possible.

One aspect of the CRC 871 is establishing a projectable disassembly procedure of the HPT blades out of the turbine disks as the initial step for regeneration. Usually, qualified and skilled workers mainly carry out the disassembly based on practical knowledge due to the difficulty of predicting process parameters [5]. However, the vague condition of turbine blades still carries the risk of further damage. That can lead not only to delays but also to parts becoming unusable for regeneration. As a result, it becomes difficult to predict

DOI: https://doi.org/10.15488/11250 publish-Ing.

and plan the disassembly time, and all subsequent regeneration processes cannot be efficiently scheduled. In order to plan and establish an automated disassembly process, it is necessary to gain information concerning parameters affecting the turbine's state. One of the main parameters is the solidification, which directly influences the disassembly force and time. In prior work, we developed a model to determine the force to disassemble aircraft engine turbine blades by simulation and experimental investigation [6][7]. Since we could not use real solidified samples of a used aircraft turbine, we create the solidified samples synthetically. The simulation and following experimental investigation so far were altogether valuable but were missing reproducibility. To face that deficit, in this work, we adapt the model with an additional variable, which represents the solidification and which we can adjust actively. Additionally, this paper aims to validate the applicability of the solidification model using replicated samples. The goal is not to exactly emulate real turbine blade disassembly but to create a basis for further experimental investigation and a method to evaluate different disassembly strategies. First, we briefly describe the existing approaches in chapter 2. In chapter 3, we will present the disassembly test rig used for our experiments, which we use to validate the solidification model described in chapter 4. Last, we will give a conclusion as well as an outlook in section 5.

2. Related Work

During operation, numerous influences change the condition of aircraft engines. The variation of flight routes, e.g. over desert or ocean, hours of operation or number of cycles, and the construction differences like material or geometry make it challenging to know in what manner the assembly connections solidify. Especially hot corrosion, caused by a mixture of components in the fuel and the intake air (e.g. sand, sea salt), is the determining factor for changing the surface's structure [8]. This not only weakens the components but also increases the effort required for disassembly. As introduced, the focus is on the disassembly of HPT blades, whose assembly connection solidify from a clearance fit during assembly to an unknown state. Process parameters like the disassembly force or time become unpredictable, which hinders the automation of disassembly. Therefore, skilled workers generally perform the disassembly manually, inter alia, using hammer strikes. However, this carries the risk of further damage.

While the automated disassembly of consumer goods like electronics [9] or electric vehicle's batteries [10] is currently widely researched, the automated disassembly of complex capital goods like aircraft engines is yet barely investigated. In order to plan and automate the disassembly of turbine blades, we plan to determine and predict disassembly parameters based on the solidification caused by wear and tear due to operation. Research on determining aircraft engine's maintenance intervals showed the possibility of predicting the remaining lifetime of turbine engine parts [11]. In prior work, we developed a model of an operationally solidified turbine-blade-disk connection to determine the disassembly parameters [6]. Using that model, we created a learning method to predict disassembly parameters based on known operational influences like hours of operation or flight routes, e.g. desert or ocean [12]. Also, we investigated decoupling the disassembly parameters from the assembly connection's geometry [7]. Using our developed technique, we intend to transcribe disassembly parameters like the force of a well-investigated geometry "A" to a yet-unknown geometry "B" without measuring it beforehand. Thus, it is no longer necessary to examine each geometry individually. Instead, we can include the geometric properties in the prediction of disassembly parameters.

2.1 Development Of A Solidification Model

Figure 1 shows the model's development, starting with an exemplary turbine disk with turbine blades inserted in a slot with a fir-tree shape in a) and a cut along the intersection line through the blade root in a turbine disk in b). Figure 1c) shows the solidification model. The solidifying force $F_s(z)$ is the force resulting from the solidification that must be exceeded by the disassembly force F_{DM} of the tool. The weight force of the blade root sample supports the disassembly force. The contact surface $A_{CS}(z)$ can be taken from CAD data and decreases during the disassembly process along the disassembly path z. It will become zero when the actual path z reaches the total disassembly length l_D. A modelled interference fit between the blade root and the turbine disk represents the solidification, where the resulting surface pressure p is the central solidifying characteristic.

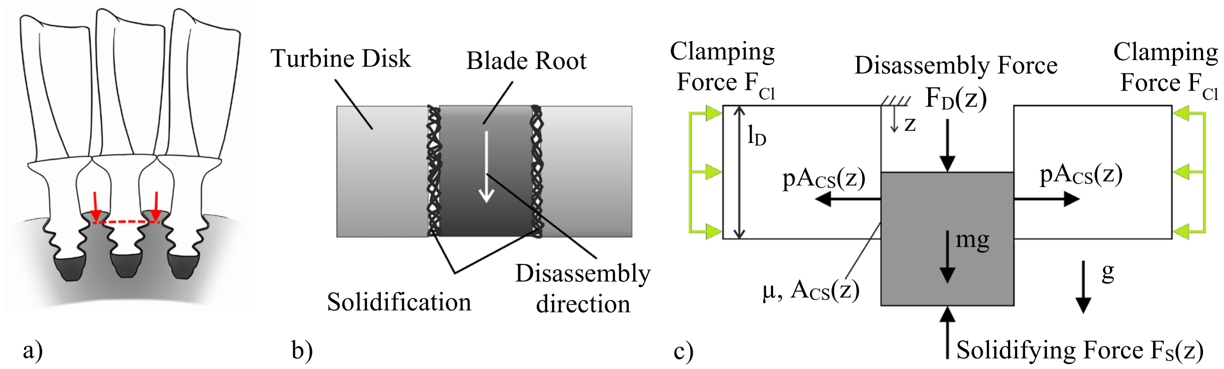


Figure 1: a) Cut through the Turbine Disk, b) Sectional view of the Turbine Disk, c) Solidification model [7]

To use the model for analysis, we transferred the model into a FEM-simulation. The simulation showed the possibility of defining and calculating constant parameters for each geometry. These geometry parameters can then be used to compare disassembly forces of two blade root shapes with a similar usage history [7]. To verify the simulation, we conducted several experiments. As aforementioned, the experimental investigation would require used aircraft turbines, which we do not have. Therefore, we had to create operational solidified disk-blade connections synthetically.

2.2 Previous Approaches For Synthetic Solidification Modeling

The first attempt was to rust samples of a blade-disk connection artificially. They consisted only of the blade root and a surrounding segment of the turbine blade. We designed three different kinds of shapes to investigate the geometry's transferability. They were similar to the fir-tree-design (Figure 1a) but with an upscaled size, consisting of an inner part, the replica blade-root, and an outer part, the replica turbine disc. The samples were made of mild steel so that they rust easily. We also specified a clearance fit so that a disassembly force in the non-solidified state is zero, just like the original connection. To artificially induce solidification, we left samples to rust in a salt spray chamber. The following disassembly test's results showed that disassembly was hardly possible. We could see wildly varying degrees of rust covering the samples. Furthermore, we could only use every probe only once for disassembly testing. Thereby the tests lack reproducibility and even feasibility. They also require a considerable amount of samples to repeat.

The second approach synthetically emulating an operational solidified blade-disk connection was to add industrial glue onto the contact surfaces of similar samples. We let it harden with varying contact pressure and times. Using an adhesive remover, we were now able to clean and therefore reuse the sample. However, compared to the rust tests, we could see much lower disassembly forces. Beyond that, we saw a very high variation in the test results. We also observed a visibly recognizable unequal glue distribution between the left and the right side and along with the shape itself. Summarily, we get a better reproducibility since we can repeat the tests more often, but we see a higher variation in the results due to the manual glue application process.

The experiments show that they lack reproducibility and are subject to variations by human inaccuracy. To tackle these obstacles, we adapt the existing solidification model. Instead of synthetically creating solidified connections, we add the clamping force F_{Cl} instead of an interference fit to induce the surface pressure caused to tackle the disadvantages and difficulties (Figure 1c). That allows us to repeat the experiments infinite times and always set the same surface pressure. A clamping unit will apply the external force. To examine and validate the adapted solidification model, we had to rework the disassembly test rig, described in the following.

3. Disassembly Test Rig

To investigate solidified turbine disk-blade root connections, we developed and built a test rig to perform disassembly tests (Figure 2). The central part is the disassembly unit, consisting of a servo motor, a gearbox, and a ball screw to produce the disassembly tool's speed. It also has a piezo-stack-actuator, a 10 kN load cell to measure the disassembly force and a pushing rod. The piezo-stack-actuator is not relevant for this paper and will be used in future work to reduce the disassembly force using micro impacts. Since the piezo actuators maximum load is 4.5 kN, the disassembly is limited to this force. The second and new element of the test rig is the clamping device to model a solidified assembly connection. The clamping unit consists of another servo motor with an integrated gearbox connected to a machine vice with a maximum force of 45 kN. Both are placed on a plain linear guidance to position the sample holder under the disassembly pushing rod manually. The sample mount is placed between the vice's jaws and includes another load cell with a maximum force of 50 kN. The samples are placed and clamped with repeatable accuracy using a locking pin and an adjusting screw. By setting the clamping force, the samples placed between the clamping unit's jaws represent a solidified turbine disk-blade connection with a reproducible degree.

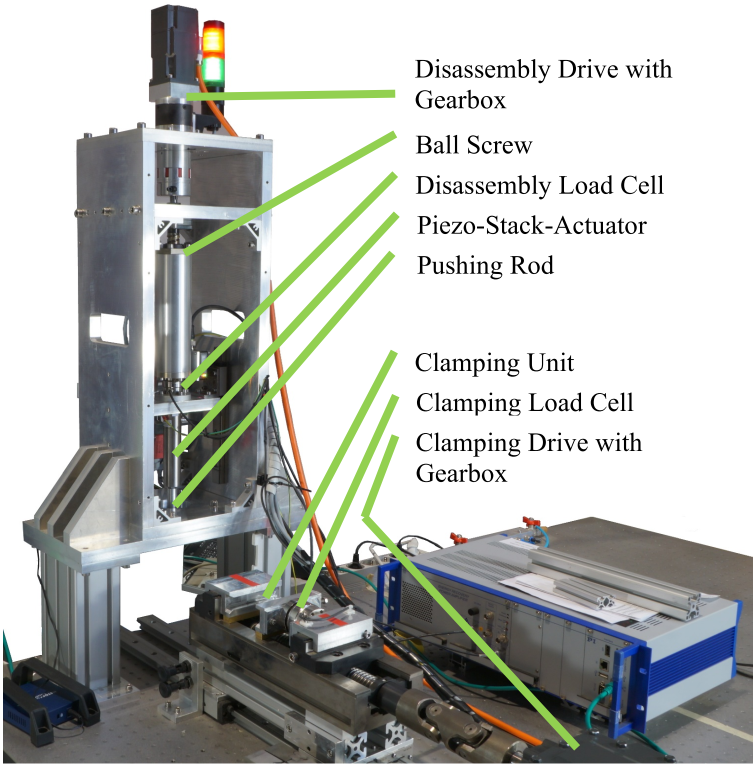


Figure 2: Turbine Blade Disassembly Test Rig

The entire system has been integrated with a Siemens PLC. With a mobile HMI panel, we can control the test rig. We use the gear transmission ratio of the gearbox and the thread pitch of the ball screw to calculate the disassembly speed from the motor's revolution speed. A laser distance sensor measures the path and determines the actual feed. To clamp the samples, we set the desired clamping force, and the servo motor turns the vice's spindle until the set value is reached. Using the PLC's internal data logging function, we can record the parameters time, disassembly speed, clamping force, distance and disassembly force. This allows us to collect the necessary data during the experiments.

4. Validation of the Solidification Model

In order to examine and validate the applicability of the solidification model, we will use the Design of Experiments (DoE), which is a technique to resolve usually complex and costly experiments efficiently. With all influences and factors included, we can obtain maximum information with a given number of experiments [13]. DoE also normalizes the data to assist the investigation and finds the values of significant factors and those with little influence. The relationship between the input variables and the effect on the output is evaluated using regression analysis, a curve-fitting method to predict an output or dependant variable based on inputs or independent variables [14]. That allows us to determine whether the variation of disassembly parameters can predict disassembly forces. It also shows the significance and amount of the influence on the output.

4.1 Full factorial DoE Analysis with Three Parameters

We will plan the experimental design by considering the main parameters for disassembly. For a profitable process from an economic point of view, the disassembly time is essential. With the relationship between disassembly path z and the disassembly time, we define the disassembly speed v_D as the first factor. The second factor is the clamping force, which resembles the solidification, as mentioned before. Furthermore, we add a third factor, the shape of the blade root, to investigate the contact surface and geometric form variation. As mentioned before, we want to decouple the disassembly parameters from geometric properties.

We considered the total samples' length of 20 mm and the disassembly time to define the disassembly speed levels of the DoE plan. Thus, we chose 1 mm/s for a slow and 10 mm/s for a fast disassembly speed to push out the blade root. In preliminary tests, we located the lower and upper level for the clamping force. When setting it lower than 2 kN, the probes repeatedly fell down before the actual test, as for higher forces than 8 kN, we were unable to push out the samples. We took the original HPT blade and redesigned two related contours to vary the blade root's shape and the contact surface (Figure 3). The samples are made of stainless steel to prevent them from rusting. Also, stainless steel is less costly than the original CSMX-4 superalloy.

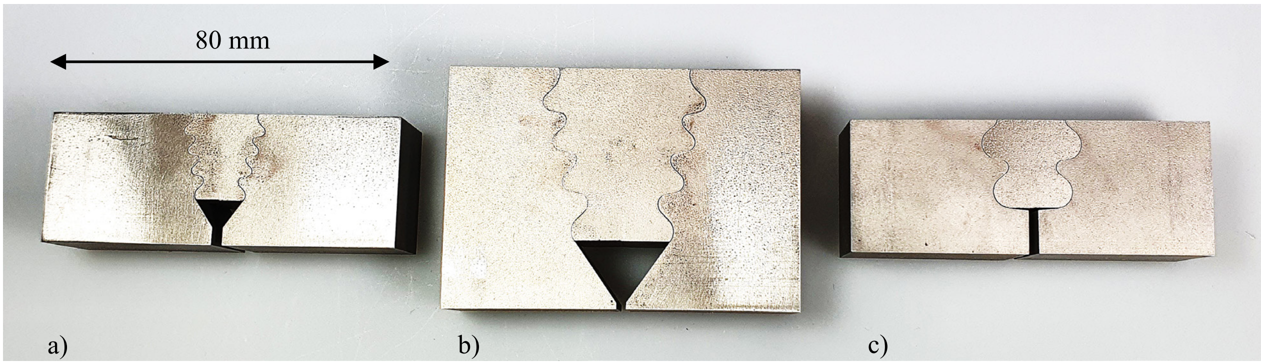


Figure 3: Overview of the different blade root shape samples: a) "Original", b) "Double", c) "B-shape"

We named the original form "Original", a contour twice the size "Double" and a third shape with the same contact surface but a different contour as the original one, "B-shape". In the experimental design, the shape of the blade root has a particular position. We divide it into three levels since it has three categories and is not a continuous variable. The levels for the factors are presented in Table 1.

Table 1: Levels of influential parameters levels in full factorial design

Factors	Low (0)	High (1)	
Disassembly Speed (v_D), mm/s	1	10	
Clamping Force (F_{Cl}), N	2000	7000	
Shape of Blade-Root (BRS), -	Original	Double	B-shape

We use the full factorial experimental design to investigate the three input variables at all possible combinations. That allows us not only to examine the influence of each factor but also the effects of interactions between the factors on the output variable. We conducted every experiment five times. The shown response variable in Table 2 is the mean over five repetitions of the disassembly force, presenting the resulting observations using the mentioned values for each factor of Table 1.

Table 2: Observation table: full factorial design

Run order	v_D	F_{Cl}	BRS	Disassembly Force (N)
1	1	1	Original	2794.51 ± 87.40
2	0	1	Original	2697.25 ± 178.30
3	1	0	Original	761.18 ± 86.63
4	0	0	Original	853.39 ± 52.72
5	1	1	Double	2362.55 ± 259.51
6	0	1	Double	2018.83 ± 207.94
7	1	0	Double	674.40 ± 96.28
8	0	0	Double	594.18 ± 54.88
9	1	1	B-shape	2200.49 ± 235.34
10	0	1	B-shape	2211.12 ± 177.74
11	1	0	B-shape	725.72 ± 122.05
12	0	0	B-shape	776.67 ± 119.38

4.2 Regression Analysis

In order to perform the regression analysis, we create a polynomial for Y, the dependant variable as a function of the explanatory variables X_i. We can also include terms to describe the influence of interdependences. To catch the diversification of the Y values, we need to add a residuum ε. With n observations, β_i the partial regression coefficients and two input variables, an exemplary equation for the regression follows to:

$$y_i = \beta_0 + \beta_1 \cdot x_{i,1} + \beta_2 \cdot x_{i,2} + \beta_3 \cdot x_{i,1} \cdot x_{i,2} + \varepsilon_i\,,\ i = 1, \dots, n \quad (1)$$

In our case, y_i stands for the estimated value for the disassembly force, and the input variables are the disassembly speed, clamping force and the blade root's shape. As mentioned before, the blade root's shape is a categorical variable. Therefore, we need two dummy or binary variables d_1 and d_2, to analyze the categories, as shown in [15]. Whenever the observation of the blade root's shape is "Double", d_1 is equal to one, otherwise zero. Similarly, we define d_2 as equal to one for "B-shape" and otherwise zero. The original shape will become the reference category since both dummy variables will be equal to zero. That results in Equation 2 for the multiple linear regression.

$$\hat{y} = \beta_0 + \beta_1 \cdot v_D + \beta_2 \cdot F_{Cl} + \beta_3 \cdot v_D \cdot F_{Cl} + \beta_4 \cdot d_1 + \beta_5 \cdot d_2 + \beta_6 \cdot v_D \cdot d_1 + \beta_7 \cdot v_D \cdot d_2 + \beta_8 \cdot F_{Cl} \cdot d_1 + \beta_9 \cdot F_{Cl} \cdot d_2 \quad (2)$$

When observing the original form, d_1 and d_2 are zero, each term containing d_1 and d_2 is dropped. Looking at the "Double" shape, d_1 will be one, and d_2 will stay zero leading to adding β_4 to β_0, β_6 to β_1 and β_8 to β_3, and dropping $\beta_5 \cdot d_2$, $\beta_7 \cdot v_D \cdot d_2$ and $\beta_9 \cdot F_{Cl} \cdot d_2$. For "B-shape", it is vice-versa.

Table 3 shows the regression's results for each input variable. The Effect-column shows the effect of changing the factors from low to high, based on each mean value. For example, increasing the disassembly speed from 1 mm/s to 10 mm/s increases the resulting disassembly force by 61.24 N. The Coef-column shows each regression coefficient β_i, and their standard errors in the SE Coef-column. They can be the same when the design matrix for calculation is orthogonal. It also contains the t-and p-values for the following goodness of fit test. The t-value is the coefficient divided by its standard error. It is used to determine the p-value from the Student's t-distribution and used for testing the significance [16].

Table 3: Coefficients table for disassembly force

Term	Effect	Coef	SE Coef	t	p
Constant		847.13	53.17	15.932	< .001
v_D	61.24	-79.67	67.26	-1.185	0.243
F_{Cl}	1649.87	1856.39	67.26	27.602	< .001
$v_D \cdot F_{Cl}$	1195.54	146.42	67.26	2.445	0.018
d_1	-215.05	-277.72	71.34	-3.893	< .001
d_2	-116.03	-39.43	71.34	-0.553	0.583
$v_D \cdot d_1$	-44.86	209.45	82.37	2.543	0.014
$v_D \cdot d_2$	-111.30	-33.32	82.37	-0.405	0.688
$F_{Cl} \cdot d_1$	761.80	-382.20	82.37	-4.640	< .001
$F_{Cl} \cdot d_2$	779.94	-483.99	82.37	-5.876	< .001

4.3 Goodness of Fit

After collecting all the data and setting up the regression model, we can now check its adequacy and fitting. The most commonly used measure for the "Goodness of Fit" of multiple regression is the multiple coefficient of determination or R^2 [16]. Its limits are between zero and one; the closer to one, the better is the fit. The calculated R^2 for our model is 0.981, as seen in Table 4. The disadvantage of R^2 is that with many input variables, its value can be high, even though one or more inputs do not affect the output. Beyond that, adding more regressors always increases but never decreases R^2 because it assumes that every predictor explains the dependant variable. The adjusted R^2 tackles this problem by considering only those input variables that affect the output. A value for the adjusted R^2 at 0.977 testifies a good fit, too.

Table 4: Regression statistics for full factorial design

Term		Value
Standard error of regression	SER	130.240
Coefficient of determination	R^2	0.981
Adjusted coefficient of determination	Adjusted R^2	0.977

After that, we will test the regression model's overall significance, i.e. test if the source of variation in Y is due to random influences or dependent on the input variables. To do this, we use null hypothesis testing. Each null hypothesis must be rejected to show a dependency between predictor and response variable. If each null hypothesis is rejected, the alternative hypothesis must be true, stating that at least one of the regressors is unequal zero.

The first assumption to investigate whether to accept or reject is the joint hypothesis. It states that all regression coefficients are zero simultaneously, as in Equation 3. We do this by the analysis of variance (ANOVA). To reject the joint hypothesis, the resulting p-value of the ANOVA-procedure must be lower than the significance level α. We chose a level of $\alpha = 0.05$, which means the risk of including effects that have no influence is 5%. A rejection of the null hypothesis will lead to accepting the alternative hypothesis (Equation 4), which states, as mention before, that at least one regressor has predictive power.

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_i = 0 \quad (3)$$

$$H_a: \beta_i \neq 0 \quad (4)$$

As seen in Table 5, the calculated p-value is lower than the significance level α, which leads to the rejection of the joint hypothesis that all input variables together have no effect on the disassembly force in the model estimated. This indicates that the regression has predictive power for the disassembly force.

Table 5: ANOVA for full factorial design

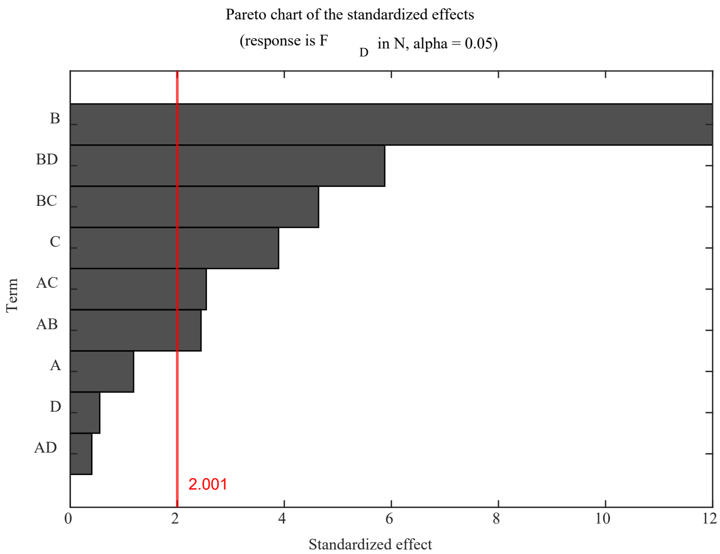
	Degrees of freedom	Sum of Squares	Mean Square	F	p
Regression	9	$4.332 \cdot 10^7$	$4.813 \cdot 10^6$	283.754	< 0.001
Residual	50	$8.481 \cdot 10^5$	$1.696 \cdot 10^4$		
Total	59	$4.417 \cdot 10^7$			

The second hypothesis testing of the regression and the model's usefulness is the significance test for each slope coefficient β_i individually. In this case, the null hypothesis states that the individual regression is not zero, as seen in Equation 5.

$$H_0: \beta_i = 0 \quad (5)$$

$$H_a: \beta_i \neq 0 \quad (6)$$

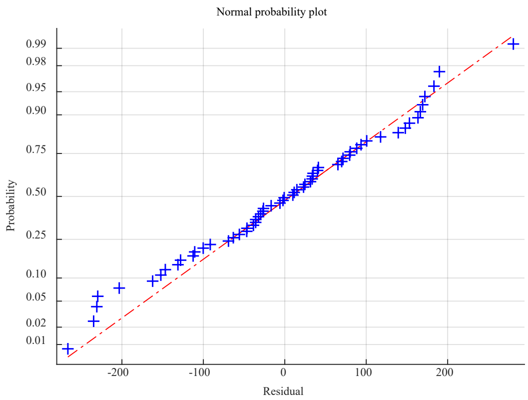
If a slope's t-value is greater than the critical t-value for the significance level alpha, the null hypothesis will be rejected. From this follows that the alternative hypothesis is fulfilled, which means that the individual regression coefficient is not zero (Equation 6). As seen in Figure 4, the clamping force and all its interactions are significant, as well as the "Double"-shape's variable and its interactions. However, the disassembly speed, the "B"-shape's dummy variable and their interaction are statistically not relevant, meaning that these terms have such a small impact that they are irrelevant for predicting the disassembly force.



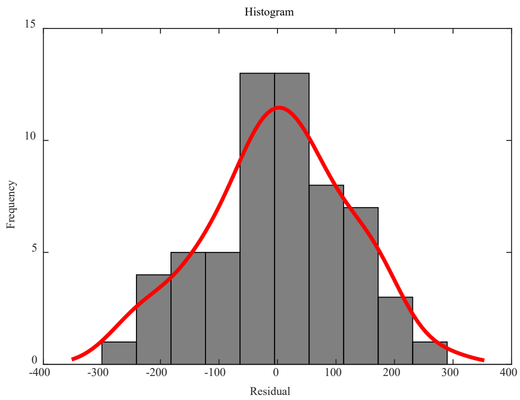
Factor	Name	Unit
A	v_D	mm/s
B	F_{Cl}	N
C	d_1	-
D	d_2	-

Figure 4: Pareto chart of full factorial design for disassembly force

Null hypothesis testing can only be applied for normally distributed estimators β_i. Since they are linear functions of each ε_i, we can test the residual's normal distribution. If the residuals are normally distributed, the estimators are as well. In our analysis, the normal probability plot (Figure 5a) indicates that the normal distribution is a good model for the results but shows a few outliers. The residuals' distribution follows a bell shape curve and also shows the presence of outliers, as they make it slightly skewed to the left and right, as seen in Figure 5b). We can therefore assume that the residuals are normally distributed, and the null hypothesis testing is valid.



a)



b)

Figure 5: Residual Plots for Disassembly Force: Full Factorial Design

The experiment shows the unexpected result that the disassembly speed has a minor influence on the disassembly force. A tenfold increase in speed from 1 mm/s to 10 mm/s decreases the disassembly force by an average of only approximately 61 N. In contrast, as expected, the clamping force has vast influence. Increasing it by 5 kN increases the resulting disassembly force by around 1.65 kN. The customized geometric shape with twice the contact surface shows a significant influence, too, in contrast to the "B-shape", whose influence can be seen as noise. According to [16], the coefficients' standard error measures its prediction accuracy. In our case, they indicate a relatively good precision. It also shows that the estimation of the input's coefficients is more precise than the dummy variable's and interactions' coefficients.

In summary, the regression analysis is suitable to describe the solidifying model. It can therefore predict the disassembly force based on the predictors speed and clamping force. Furthermore, we can include the geometric shape in the prediction model. The exact dependence of the geometric shape on the output variable and a decoupling from geometric properties will be investigated in future work. Also, we need to determine the correlation between the operational influences (flight route, number of cycles) and the clamping force.

5. Conclusion and Outlook

This work's objective is to adapt the existing solidification model and to validate its applicability. As mentioned at the very beginning, our previous research's main issue was the lack of reproducibility. We showed that we could model a solidified turbine disk-blade joint in a reproducible manner by an additional force using a clamping unit. The experiment was planned using DoE to gather data. With setting up a regression, we created a basis to predict the disassembly force to plan disassembly tools and time based on performed experiments and gathered data. We can also automated disassemble simple turbine blade samples using the disassembly test rig and measure the necessary information to predict the disassembly force.

In future work, we will add the presented piezo-stack-actuator to confirm the possibility to decrease the needed disassembly force, as seen in past investigations. Since the manufacturers instruct specifications and damage patterns for accepting or rejecting parts for reuse, we need to reduce the pressure on the surface as much as possible. Together with the prediction model, we can establish a foundation for an adaptive and component friendly disassembly process. Additionally, we need to investigate how operational parameters like the number of cycles and flight routes affect the modelled solidification. Eventually, we want to use the developed knowledge to predict the disassembly parameters depending on the operation. That will support aviation MRO planners and companies to plan an automated and component friendly disassembly efficiently.

Acknowledgements

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 871/3 – 119193472

References

[1] Lambert, A.J.D., 2003. Disassembly sequencing: A survey. International Journal of Production Research 41 (16), 3721–3759.

[2] Das, S.K., Naik, S., 2002. Process planning for product disassembly, International Journal of Production Research Vol. 40:6, pp. 1335-1355.

[3] Rupp, O., 2001. Instandhaltungskosten bei zivilen Strahltriebwerken, Deutscher Luft- und Raumfahrtkongress 2001, DGLR-2001-008, Hamburg

[4] Ackert, S., 2011. Engine Maintenance Concepts for Financiers, Technical Report, Aircraft Monitor

[5] Tao, W. Huapeng, D., Jie, T., Hao, Wa., 2015. Recent Repair Technology for Aero-Engine Blades, Recent Patents on Engineering Vol. 9, No 2, pp. 132-141.

[6] Wolff, J., Yan, M., Schultz, M., Raatz, A., 2016. Reduction of Disassembly Forces for Detaching Components with Solidified Assembly Connections, Procedia CIRP 44, pp. 328-333.

[7] Wolff, J., Kolditz, T., Fei, Y., Raatz, A., 2018. Simulation-Based Determination of Disassembly Forces, Procedia CIRP 76, pp. 13-18.

[8] Betz, W., Huff, H., Track, W., 1976. Zur Bewertung von Schutzschichten gegen Heißgaskorrosion an Gasturbinenschaufeln. Mat.-wiss. u. Werkstofftech. 7 (5), 161–166.

[9] Vongbunyong, S., Kara, S., Pagnucco, M., 2013. Application of cognitive robotics in disassembly of products. CIRP Annals 62 (1), 31–34.

[10] Wegener, K., Chen, W.H., Dietrich, F., Dröder, K., Kara, S., 2015. Robot Assisted Disassembly for the Recycling of Electric Vehicle Batteries. Procedia CIRP 29, 716–721.

[11] Müller, M., 2013. Untersuchungen zum Einfluss der Betriebsbedingungen auf die Schädigung und Instandhaltung von Turboluftstrahltriebwerken, Institut für Luftfahrtantriebe, Universität Stuttgart

[12] Wolff, J., Kolditz, T., Raatz, A., 2019. A Learning Method for Automated Disassembly, Precision Assembly in the Digital Age, IPAS 2018, Springer, Cham, vol 530, pp. 63-71.

[13] Guthrie, W.F., 2020. NIST/SEMATECH e-Handbook of Statistical Methods (NIST Handbook 151)

[14] Toutenberg, H., 2006. Arbeitsbuch zur deskriptiven und induktiven Statistik, Springer, Berlin, pp. 227-239.

[15] Allen, T. T., 2019. Introduction to Engineering Statistics and Lean Sigma, Second Edition, Springer, London, pp. 397-401.

[16] Gujarati, D. N., Porter, D. C., 2008. Basic Econometrics, Fifth Edition, Irwin/McGraw-Hill, New York

Biography

Richard Bluemel (*1987) studied mechatronics engineering at Leibniz University Hannover. He was employed at International Automotive Components Group (IAC) as a construction engineer in the field of special mechanical engineering (2014-2019) before returning to the Leibniz University Hannover as a research associate at the Institute for Assembly Technology (match).

Annika Raatz (*1971) has been the head of the Institute of Assembly Technology (match) since its establishment in 2013. Prof. Dr.-Ing. Annika Raatz is a member of the Society for Assembly, Handling and Industrial Robotics (MHI) and the Academic Association for Production Technology (WGP). Her main research topics are robot-aided handling processes, machine concepts for handling, assembly and production automation and soft material robotic systems.

2nd Conference on Production Systems and Logistics

Consideration Of RFID Systems In The Factory Planning Process

Christian Looschen[1], Simon Berger[2] , Miroslav Lach[3], Johannes Fottner[1]
[1]Chair of Materials Handling, Material Flow, Logistics , Technical University of Munich, Garching, Germany
[2]Fraunhofer Institute for Casting, Composite and Processing Technology IGCV, Augsburg, Germany
[3]Associate Professorship of Microwave Engineering, Technical University of Munich, Munich, Germany

Abstract

RFID technology enables products and load units to be identified in the absence of a direct line of sight. In recent years, it has grown into one of the principal tools of identification used in industrial environments. For it to function properly, efficiently and reliably in such locations, factory planners require comprehensive integration of an RFID system. Although planning methods and tools for integrating RFID into existing production and logistical environments are available, it takes a great deal of time and expense to implement the procedures successfully. This is because logistics and production systems and processes need to be adapted to incorporate an RFID process, which discourages many companies from installing it in their production. Furthermore, to avoid identification problems in production and material flow processes, factory planners need substantial knowledge of the parameters that influence RFID.

This paper presents a generic planning method which considers the conceptual design and implementation of RFID systems in an industrial environment from the beginning of the factory planning process. The aim of the method is to give factory planners and RFID experts a guideline on how to plan and implement RFID, by taking an interdisciplinary approach. After first reviewing current RFID planning methods, the paper presents a new six-phase planning method the aim of which is to enable experts in the industry to plan their production, logistics and RFID systems simultaneously, with due regard for the individual requirements of each system.

Keywords

RFID Systems; RFID Planning; Factory Planning; Planning Processes

1. Introduction

RFID is a driving technology that enables processing companies to identify tagged products in the absence of a line of sight and to seamlessly track the goods at a production site. RFID tags are applied to goods and read by a dedicated RFID reader. The advantage of RFID technology is that it enables production and logistics to be controlled effectively and gives management an overview of material flows. Although RFID has further advantages, implementing this technology in a processing site is a challenging task. RFID has many restrictions, which managers and production site planners need to be aware of [1]. For example, metallic objects can interfere with or even block the electromagnetic waves of the RFID systems, which can lead to a breakdown in communication [2]. Other problems in the RFID system during the production process include RFID tags that cannot be read or electromagnetic waves from different RFID readers that interfere with each other.

When such problems occur, factory managers have to make adjustments not only to the RFID technology itself but also to the production site and its logistics. This can seriously disrupt the production process, which in turn leads to a loss in profits. To prevent such problems, factory planners have to consider RFID when designing a new factory, redesigning an existing one, or implementing RFID for the first time at an existing site. To achieve this, RFID must be an integral component of the factory planning process. This concerns not only the planners and their methods, but also extends to their tools and way of thinking. An RFID system simulation may be an appropriate tool for such a purpose.

Many tools are already available that simulate RFID and the propagation of electromagnetic waves. Furthermore, manufacturers of RFID technology already possess detailed knowledge of the limitations and requirements of their products. Manufacturers use a variety of computer-aided models (CAX), which RFID experts can employ for their tools. These tools already exist and set the stage for planners of logistics, production and RFID systems to design factories in an integrative manner. However, it is necessary to address not only the questions of how and when different planners will work together but also how to implement the data flows of the different data formats between the planners and how the RFID simulation tools will be used to plan a factory. [3]

The authors of this paper present a method by which planners can design and lay out logistics, production and RFID systems for planning factories in an integrative manner.

2. Current approaches to planning RFID systems

Several RFID planning processes have been developed over the last few years, most of them concentrating on integrating the technology in logistical systems. The common ground of these processes is that they focus on integrating RFID into existing systems. Furthermore, the procedures in most of these models are generally the same:

1. Define the aims of using RFID;
2. Analyse the place of operation;
3. Derive the system requirements and boundary conditions;
4. Develop an application concept;
5. Implement the established concept.

Several good examples of these methods have been published [see 4-9].

Hellström proposes an RFID implementation process based on an IT implementation model [4]. The model is derived from data gathered from two use cases and divides the RFID implementation into seven stages: initiation, adoption, adaptation, acceptance, routinisation and infusion. The author points out that although all the stages can be performed simultaneously, if one step proves inconclusive regarding the benefits or implementation possibilities of RFID, it would be necessary to iterate previous steps. The activities of each implementation stage are outlined in the model. However, the model does not go into detail about how the activities should be conducted.

The method by *Donath*, as described in [5], introduces a process for implementing RFID in small and medium-sized enterprises (SMEs) and proposes an implementation in six phases rather than seven as in [4]. The author focuses on pointing out at which time of each phase different tests and analyses are performed to verify the concept for integrating RFID in existing logistical process. Among other things, the author proposes economic analyses, components, integration processes, systems and acceptance tests. The tests are categorised as laboratory (or synthetic) or on-site-tests. Again, the procedure is not further elaborated.

A similar method was developed by *Fruth* in [6], who also proposed an RFID implementation process for SMEs. Dividing the process into five phases (definition, concept, objective design, pilot, and implementation), the authors place the activities relating to implementation into one of three categories: performance processes, supporting processes or management processes. Each category of activity thus plays a different key role.

A different type of process model is presented by *Vogeler* in [7]. Based on case studies conducted in the textile industry, the author presents a model consisting of a series of flexible iterative steps, which are categorised into four activities: controlling, security, design and support. The steps are not necessarily followed in a linear fashion but can be repeated in one or more phases of the implementation process. A process intensity diagram shows where the planner of the RFID system should initiate each of the planning steps. The diagram also suggests milestones for each activity, but these may be moved depending on the project in which the model is used.

Fischer develops a modular concept for planning RFID-supported logistical processes in [8]. Instead of investigating the various planning phases, the author considers the different RFID modules. These are the RFID technology, the object of identification, and the identification process. In the first step, the requirements, boundary conditions, resources needed, and objects of interest are defined for each module. This is followed either by an evaluation of existing solutions for use in the current tasks or the development of new solutions.

A different approach to those given above is presented in [9] by *Steinhaus*. Instead of defining specific tasks in a specific order, the author presents questions intended to help SMEs implement RFID in their businesses. The questions are grouped into preparation, organisation, technology and process management. They are published in a guideline, which also gives examples of how RFID can be implemented in specific logistical processes.

3. Integrated RFID planning method

3.1 Method development

As stated in the previous section, current RFID planning methods focus on the implementation of RFID systems in an existing production or logistical environment. However, no link exists between the factory activities and the work of RFID planners. This section presents an integrated planning process, in which the two interacting aspects of RFID and factory planning are considered simultaneously from the beginning of the factory project. The aim of the method is to provide planners with a holistic and interdisciplinary guide, which they can use in a generic factory planning project. The method was mainly devised and evaluated for green field planning.

The method is based on the processes described in the Munich Model of Methods by *Braun & Lindemann* [10]. First, the existing methods, which were introduced in Section 2, were analysed. It was determined that since the respective planning procedures and phases in RFID are similar to those in the factory environment, they can be superimposed. The method developed is founded on the planning processes in [5,8,6]. However, as can be seen from Figure 1, the phases of the different processes do not coincide entirely.

For this reason, the RFID planning processes were broken down and the individual steps reallocated into the factory planning processes from [11], which is a more detailed extension of the standardised factory planning process of [12]. Further steps were added to meet the requirements of an integrated planning process. These steps mainly concerned the estimation of signal coverage using computational electromagnetics (CEM) software as an additional planning tool. The method and its phases are presented schematically in Figure 1. The different colours inside the phases correspond to the qualitative integration of the RFID planning steps of the respective authors into the factory planning process in [11]. Dark red indicates the additional planning steps devised for this new method, which include the use of CEM.

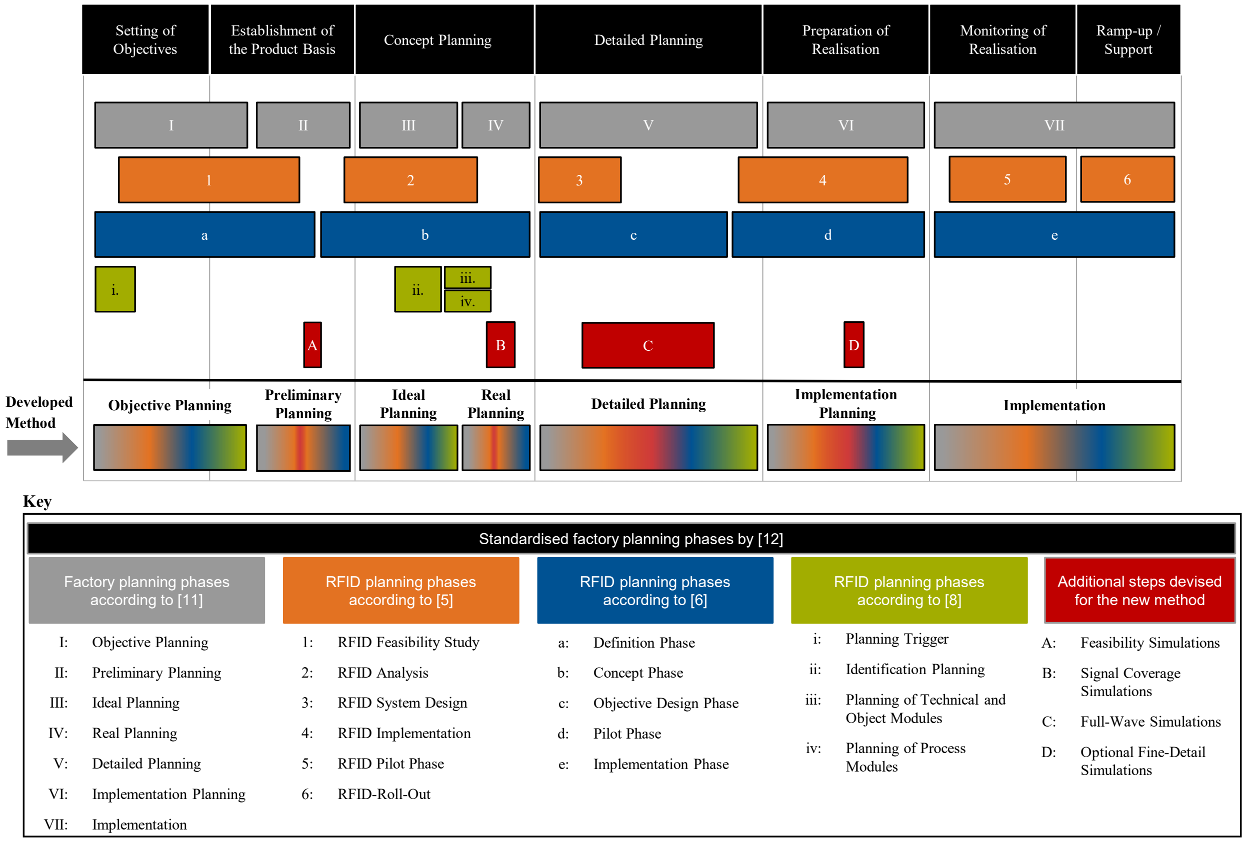


Figure 1: Qualitative comparison of the planning phases in [5], [6], [8], [11] and the method with the additional planning steps described section 3.2 in reference to the phases of the standardised factory planning process of [12].

The following section presents an overview of these steps in the different phases of the integrated planning process.

3.2 Planning phases of the integrated RFID and factory planning process

3.2.1 Objective planning

The integrated RFID and factory planning process commences with an objective planning phase, in which the objectives of the factory building project are set and the basis of the planning process established. The steps included in the general factory planning process phase determine the envisaged product range, set key dates, work packages and planning milestones, and compile an initial financial framework for the project as a whole. Also included in the first steps of the objective planning phase is the establishment of the data sets required for the subsequent planning process. The data can be obtained by revising past planning projects and product information, researching the latest government regulations, and gathering information on the state of the art in production as well as in logistics technology and processes.

The integration of RFID planning in this phase extends the general factory planning phase by ensuring that RFID is considered in the first few steps of the project. This includes analysing the use of RFID in past projects or researching the state of the art of different RFID systems for similar processes. Additionally, factory planners need to identify initial potential areas of deployment for RFID systems in the planned factory environment.

At the end of this phase, planners should have a portfolio of information gathered during the phase, as well as a complete project plan containing work packages and schedules for the project as a whole.

3.2.2 Preliminary planning

Planners use the data gathered in the objective planning phase to build an initial theoretical concept of the production and logistical systems and processes in the preliminary planning phase. The concepts in this phase do not include a factory layout but concentrate on the procedure and sequences of different processes in the production and logistical environment, and assemble them into flow charts, functional diagrams, etc.

The requirements of the RFID systems to be used in the production and material flow processes can then be derived from the resulting production program and production & material flow chart. This involves a requirements analysis, which includes analysing possible tag positions on the product, as well as tag types, reading distances, reading rates, and material influences. The requirements analysis should not only focus on the current production plans but also on planned future production projects and possible uses of RFID in the company, thus taking into account the adaptability of the factory. It is crucial that RFID system planners work closely with the planners of the logistics and production systems. Whenever there are any changes in the concept of any subsequent systems, these can affect the requirements which the RFID systems have to fulfil.

Once the preliminary requirements are available to the RFID planners, the next step requires them to set up initial concepts for the application of RFID tags and to perform the initial tests required for a feasibility study of these concepts. These can include synthetic laboratory tests (such as described in [13] and [14]) for example or simulations (e.g. as presented in [15]) to show how RFID tags can be applied to assembly components of product or loading units and how different setups (e.g. tag position, tag type, antenna type, etc) can affect system efficiency. The aim of these tests is to validate the application concepts developed and to determine the range in which RFID can be used in the various production and logistic processes.

Once initial production, material flow and RFID concepts are available, a rough hardware and software demand and cost assessment is conducted and compared to the financial framework setup from the objective planning phase. If these coincide, the management can accept the concepts, and the manufacturing and material flow processes can be adopted. In the next step, the factory planners can continue planning the ideal factory layout.

3.2.3 Ideal planning

The ideal layout planning phase is the first part of the factory planning process and culminates in the presentation of an initial layout design at the end of the phase. Schematics for the workflow in the factory can be derived from the production program and the manufacturing and logistical process concepts. By analysing these, planners can build up different scaled 2D floor plans, which include dimensions of factorial subsystems (such as the areas of different manufacturing and assembly zones) and the visualisation of the material flow between these areas.

At the same time, RFID experts can use the workflow schematics to specify the processes in which identification operations are to be used. It is important that there is an exchange of information between factory and RFID planners at this stage of the project, as further layout designs need to be adapted to the RFID concepts (such as space, cable planning, information flow planning, etc.) and vice versa. Planners should always bear in mind that the factory concept can be revised more easily at early stages. The cost and effort of changes rises in later stages.

Further analysis regarding the arrangement and implementation of RFID in factory processes requires more specific configuration of the various factory subsystems. This is why no additional RFID steps have been added to this phase of the integrated planning process. However, further steps will be conducted in the next phase, in which the real layout is planned.

3.2.4 Real planning

The aim of the real-layout planning phase is to concretise the ideal layout plan from the previous phase in an area-specific presentation of the factory area, or real layout. In this depiction of the factory, the areas for the machine groups, workstations, material flow paths etc. are arranged inside the areal-scaled layout plans of the ideal layout.

Up until this phase, planners had been devising RFID application concepts in the factory operations using workflow schematics created by factory planners in the initial phases of the planning process. In the next step, the planners will be designing the first factory plans with different real layout variants. This allows RFID experts to concretise the arrangements of the RFID systems inside the factory and identify specific boundary conditions which planners need to bear in mind when generating their RFID implementation concepts. Concretising the arrangement of RFID systems means specifying the number of identification points (points in a production or logistics process where an RFID tag may be read) and arranging them throughout the factory layout. In addition, a detailed installation approach for the RFID antennas has to be created, based on the experimental results of the RFID feasibility study in the preliminary planning phase (i.e. positioning the tag on the product or loading unit). It is also necessary at this point to consider how and how often identifiable products will pass by the identification point, as well as how and where RFID tags will be placed on these components. This also necessitates further simulation studies (e.g. using ray-tracing simulations as in Figure 2 and described by *Bosselmann* in [16]) to verify the layout of the RFID concept and the signal coverage within the vicinity of the antenna. Criteria for verification include the locations of critical wave dead spots and overreach. Dead spots, where tags cannot be read, have to be ruled out to ensure high system reliability, while overreach results in tags being identified outside the designated reading area.

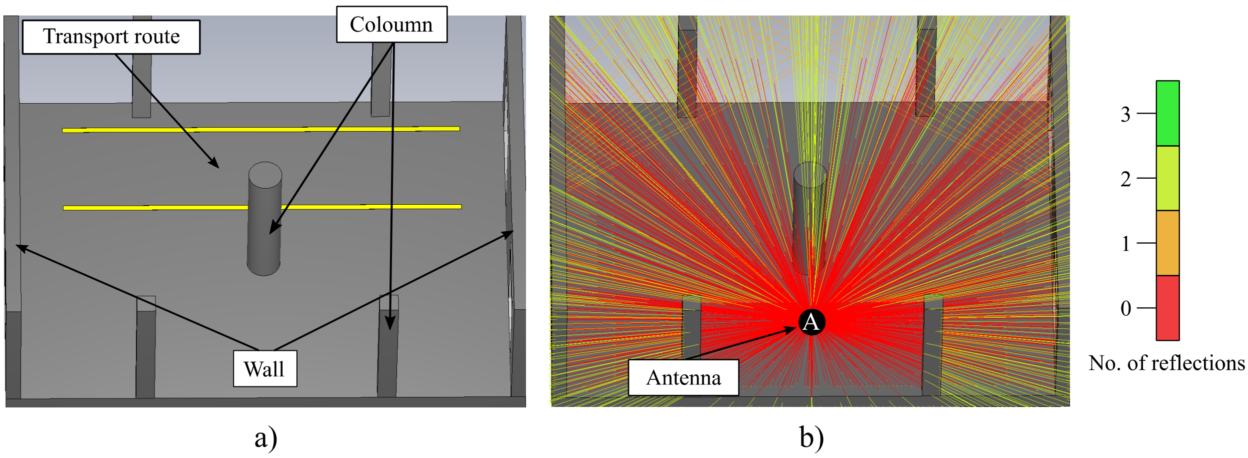


Figure 2: An example 3D model of an empty factory hall (a) and the same hall with an RFID antenna (b). Figure (b) also shows an analysis of indoor wave propagation utilizing ray-tracing simulations. (A) indicates the RFID reader antenna. The rays represent an approximation of the field distribution of the magnetic waves inside the hall.

These studies require 3D models of the factory environment, in which computer models of the RFID systems can be integrated. Also, these models should ideally be used to arrange the factory systems in the real layout. To shorten simulation times, it is advisable to divide the factory model into several smaller sections. Only those sections relevant to the RFID process are then analysed in the simulation studies. It should be noted that analysis of different sections can lead to different results if sections are combined or the area of a section increases. Again, factory and RFID system planners have to work together, and, where necessary, implement changes into their concepts to ensure that both systems work optimally.

While analysing the arrangement of the RFID systems, planners also need to draft the information flow system between the RFID and the other factory entities, in a parallel process.

Once different arrangement concepts for the factory and RFID systems have been completed and combined into a number of real layouts, they have to be assessed and the preferred alternative chosen. Relevant criteria to be considered for assessment are functionality, costs, profitability and goals set at the start of the factory planning process.

3.2.5 Detailed planning

The final step is the detailed planning phase, and it is here that the layout of the factory is detailed in full. By the end of this phase, planners will have developed a layout plan that can be used to build the factory in the subsequent implementation phase. The plan includes dimensionally accurate factory areas, the positions of workspaces, machines, equipment, and material handling and storage systems. It also includes installation details for IT and communications equipment.

The most detailed plan of the factory systems is used to finalise the RFID implementation concept and conclude the arrangement of the systems, the choice of hardware and software, and the planning and execution of the final feasibility study.

The feasibility study normally consists of hardware and process tests conducted in a realistic environment. We have extended the study by incorporating wave simulations in the test procedure, a process not included in previous methods. These simulations can be performed in parallel with the real-world tests. Both types of experiment need to be prepared simultaneously so that they can analyse the same settings. The objectives of the experiments are derived from the project goals and RFID requirements, and the setup is determined from the arrangement of the factory and RFID systems. The quantity and duration of the experiments to be conducted are determined by building experimental plans using design of experiments (DoE) processes, as described in [17,18].

To use simulation as an analysis tool, further steps are required to ensure that the simulation is as realistic as necessary. Full-wave simulations are recommended for this purpose, example results of which are shown in Figure 3. These are more accurate than ray-tracing simulations in predicting the field distribution and power of electromagnetic waves, but they require a greater computational effort. It must first be ensured that the correct material parameters have been assigned to the objects in the simulation model and that the correct boundary conditions for the model are set. Furthermore, a simulation method for depicting wave propagation has to be chosen and the 3D model adjusted accordingly. Among other things, this means simplifying 3D models of machines to shorten the simulation times.

Once both test types have been completed, the results are compared and rated in relation to the project's goals and RFID system requirements. The concepts may also have to be adapted as necessary. If all requirements are fulfilled and the concepts for the RFID and factory systems coincide, a final evaluation (functionality, cost efficiency, etc.) of the factory and its concept as a whole is conducted before management approval of the plan is given in the final two phases.

3.2.6 Implementation planning and implementation

The last two phases of the factory planning process are concerned with the planning and execution of the factory implementation. This includes the call for tender for the different systems and the planning of system assembly and installation (including RFID) as well as human resources.

Once the exact models of the machine and RFID systems are available, last-minute wave propagation simulations can be conducted if deemed necessary by the planners. These simulations are as accurate as possible, since all boundary conditions, material parameters and machine dimensions used in the factory are fully defined by this stage. The planner can use these to make last minute fine adjustments to tag or antenna positions at the various identification points, and, if required, increase the tag read rate and probability, and in turn, the efficiency of the RFID system as a whole.

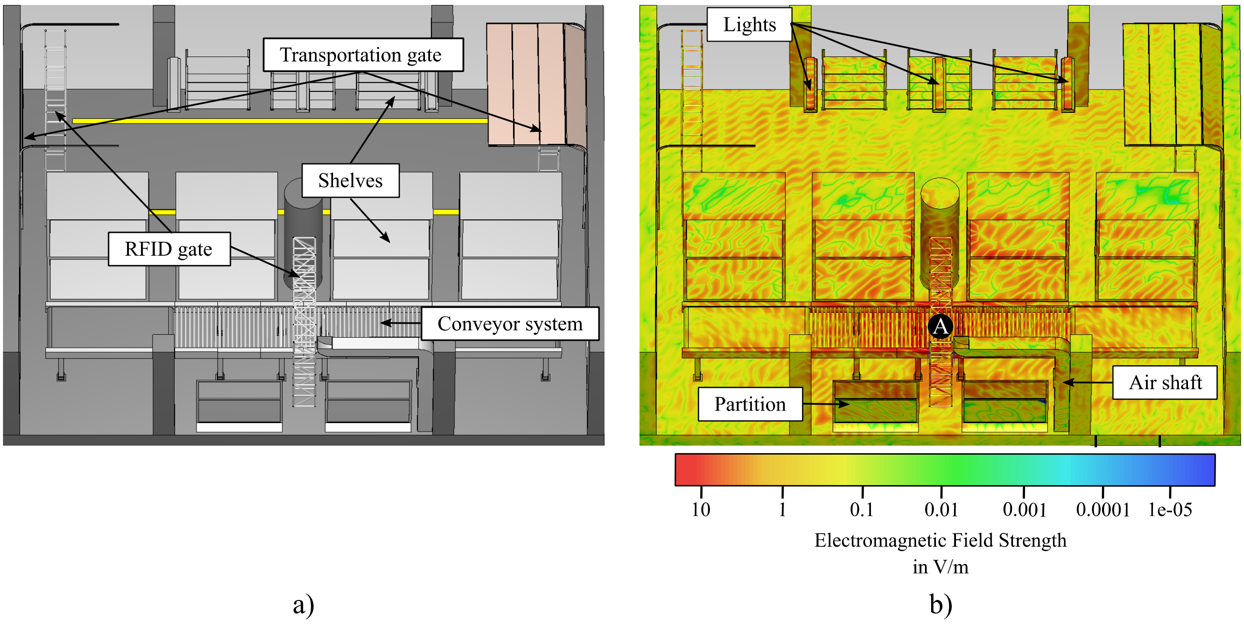


Figure 3: The factory hall from Figure 2 with logistical systems (a) and the same hall with an RFID antenna, whose wave propagation is analysed by electromagnetic full-wave simulation methods (b). The latter shows the electromagnetic field distribution emitted by the antenna (A) in V/m throughout the hall.

Once a time schedule has been set for the building and installation project, the construction of the factory can begin. After completion of the factory, the project conclusion process begins, which includes calculating the total costs of the project and completing the project documentation. Pilot production, which takes place simultaneously, entails final process and system testing and also includes any last-minute changes.

Once the pilot production has been completed and the results accepted by the management, serial production can commence, signalling the conclusion of the integrated factory planning process.

4. Summary

This paper has shown that although planning processes already exist for integrating RFID into logistics and production systems, these models only consider the integration of RFID systems in existing factory systems. As this can lead to difficulties with refurbishment processes or even to production stops, the use of RFID needs to be considered from the very start of the factory planning process. The integrated planning process presented in this paper proposes steps to achieve this by combining the existing RFID planning processes with classic factory planning methods and extending these steps where necessary.

After briefly describing the development of the method, the authors elaborate the RFID planning steps that extend the factory planning process. The six stages of the factory planning process considered are the objective planning, preliminary planning, ideal planning, real planning, detailed planning and implementation planning phases and the implementation phase. The presented planning model shows that RFID can be considered in every phase of the factory planning process, and that, like the factory environment, the level of detail of the RFID system concept grows over the course of the planning process. Furthermore, additional analysis tools for evaluating the RFID system concepts are included in the various planning phases. Ray-tracing and full-wave propagation simulations are examples of such tools. The model also demonstrates that simultaneous, interdisciplinary planning of the factory including RFID is necessary, as the different systems (i.e. their dimensions and arrangements) affect each other. As it is far more expensive in terms of resources to make changes at a later phase of a planning project, it is crucial that work is conducted in an interdisciplinary manner from the very start of the project.

The method presented demonstrates that every phase of the factory planning process contains steps in which RFID systems can be considered. These have an impact on the different production and logistic systems and their processes, and conversely, the layout and processes of the system also have an impact on the composition of the RFID systems as well as on the process in which RFID is to be integrated. This is one reason why RFID should be taken into consideration right from the beginning of a factory planning project. Furthermore, retrospective integration of RFID into existing systems is more costly in terms of resources and replanning activities.

The next step is to verify and evaluate the functionality of the method introduced in this paper. Moreover, further work is needed to detail the interchange of data produced by the production, logistics and RFID planners and how each planning department can use the data of the other departments.

Acknowledgements

The findings presented in this paper were achieved in the "IntegRFID" research project, which is funded by the German Research Foundation (DFG) under the project number 426221637. We would like to extend our thanks to the DFG for supporting our work.

References

[1] Bekkali, A., Zou, S., Kadri, A., Crisp, M., Penty, R., 2015. Performance Analysis of Passive UHF RFID Systems Under Cascaded Fading Channels and Interference Effects. IEEE Trans. Wireless Commun. 14 (3), 1421–1433.

[2] Schneider, O., Hohenstein, F., Günthner, W.A., 2011. Bewertung von Methoden hinsichtlich einer ganzheitlichen Prozessdarstellung, 18 pp.

[3] Reinhart, G., 2007. Einführungsstrategien für RFID in produzierenden Unternehmen. Seminarberichte 89: RFID in der Produktion. iwb, Technische Universität München.

[4] Hellström, D., 2009. The cost and process of implementing RFID technology to manage and control returnable transport items. International Journal of Logistics Research and Applications 12 (1), 1–21.

[5] Donath, S. (Ed.), 2010. Implementierung von RFID Lösungen in Unternehmen, 6 pp.

[6] Fruth, A.J., Günthner, W.A., Krcmar, H., 2012. Methodik für die Durchführung von RFID-Implementierungsprojekten in der Logistik bei mittelständischen Unternehmen. Zugl.: München, Techn. Univ., Diss., 2012. Universitätsbibliothek der TU München; Techn. Univ. fml – Lehrstuhl für Fördertechnik Materialfluss Logistik, München, 20 pp.

[7] Vogeler, S., 2009. Entwicklung eines Vorgehensmodells zur Implementierung der RFID-Technologie in logistischen Systemen am Beispiel der Bekleidungswirtschaft. Dissertation.

[8] Fischer, R.J., 2016. Modularisierung als Planungsansatz für RFID-Anwendungen in der Intralogistik. fml - Lehrstuhl für Fördertechnik Materialfluss Logistik, München, 249 pp.

[9] Steinhaus, A., Hütter, S., 2013. Leitfaden zur Implementierung von RFID in KMUs, Saarbrücken, 52 pp.

[10] Braun, T., Lindemann, U. (Eds.), 2003. Supporting the Selection, Adaptation and Application OF Methods in Product Development, Stockholm, 10 pp.

[11] Grundig, C.-G., 2018. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 6., neu bearbeitete Auflage ed. Hanser, München, 333 Seiten.

[12] Verein Deutscher Ingenieure, 2011. VDI 5200 - Blatt 1 - Fabrikplanung Planungsvorgehen. Beuth Verlag, Berlin, 24 pp.

[13] GS1, 2017. Tagged-Item Performance Protocol (TIPP) Tagged Item Grading: Testing Methodology Guideline: Provides the testing procedures for tagged-item grading, 17 pp.

[14] Günther, W.A., Atz, T., Klaubert, C., Salfer, M., 2011. Forschungsbericht RFID-MachLog: Methodik für UHF-RFID-Machbarkeitsstudien.

[15] M.M. Keskilammi, L.T. Sydanheimo, M.A. Kivikoski, 2004. Modeling Electromagnetic Wave Propagation in Paper Reel for UHF RFID System Development. Proceedings of the 5th International Conference on Machine Automation ICMA2004, November 24-26 2004, Osaka, Japan, 565–568.

[16] Bosselmann, P., Rembold, B., 2006. Investigations on UHF RFID wave propagation using a Ray Tracing Simulator. Frequenz : journal of RF/microwave engineering, photonics and communications (60), 38–45.

[17] Antony, J., 2003. Design of experiments for engineers and scientists. Butterworth-Heinemann, Oxford.

[18] Siebertz, K., van Bebber, D., Hochkirchen, T., 2010. Statistische Versuchsplanung: Design of Experiments (DoE). Springer Berlin Heidelberg, Berlin, Heidelberg, Online-Ressource.

Biography

Christian Looschen is a research assistant at the Chair of Materials Handling, Material Flow, Logistics at the Technical University of Munich (TUM) since January 2020. His field of research is the planning of RFID in intralogistical systems. He graduated from the Faculty for Mechanical Engineering of the TUM in 2019.

Simon Berger is a research assistant at the Fraunhofer Institute for Casting, Composite and Processing Technology. His research focuses on identification systems, sensor systems and robotics. He has contributed to and managed several projects in these fields. He was nominated for the "Deutscher Zukunftspreis 2018" in the field of underwater robotics in the context of nuclear decommissioning.

Miroslav Lach is a research assistant at the Associate Professorship of Microwave Engineering at the Technical University of Munich (TUM) since October 2019. He received his M.Sc. degree from the Department of Electrical and Computer Engineering at the TUM in 2019. His field of research is the analysis and simulation of UHF-RFID systems and their reciprocal effects on the operational environment.

Johannes Fottner received a Dr.-Ing. degree in mechanical engineering from the Technical University of Munich (TUM), Munich, Germany in 2002. Since 2016, he has been the head of the Chair of Materials Handling, Material Flow, Logistics at the TUM. His current research interests include investigating innovative technical solutions and system approaches for optimizing logistical processes.

2nd Conference on Production Systems and Logistics

A Review Of Data-Based Methods For The Development Of An Adaptive Engineering Change System For Automotive Body Shop

Günther Schuh[1], Georg Bergweiler[1], Falko Fiedler[1], Viktor Slawik[1], Christian Ahues[1]

[1]*Werkzeugmaschinenlabor WZL der RWTH Aachen University, Aachen, Germany*

Abstract

The automotive body shop with its high investment costs and resource intensive high automation level is not sufficiently responsive to the demands of agility and adaptivity. The coexistence of combustion engine and e-mobility platforms, but also design modifications on products during production phase, require constant and agile product adaptation in automotive structural components.

These changes to the structural car body design have different effects on production planning. However, in some cases they have a considerable influence on the equipment design, e.g., the choice of jig, fixture, and joining technology and the number of robots used. This means that investment decisions for the procurement of new equipment or even entire production systems are also determined. In particular, late introduced new requirements or engineering changes in the product development process (PDP) have a massive impact on the quality of planning and compliance with established premises at the beginning of the PDP. Measures for the holistic consideration of engineering changes and their influences require both a product-side and a production-side review of influences.

In this review paper, existing approaches for the product and production-side handling of requirements and engineering changes are analyzed by a systematic literature review. Existing data-based methods are analyzed with regard to their applicability on engineering change management (ECM) in body shop. Based on this, enabling methods are assessed concerning their suitability for the transfer to a holistic, adaptive, and data-based ECM system for the PDP of automotive car bodies.

Keywords

Requirements Management; Engineering Change Management; Change Request; Artificial Intelligence; Vehicle Production; Versatile Production; Production Intelligence

1. Motivation

Shorter product life cycles and a dynamic, interdisciplinary product development process (PDP) characterize automotive body design [1]. In fast-moving agile development of product and production, requirements management (RM) and engineering change management (ECM) are therefore becoming increasingly important. The rise of e-mobility and the associated profound alterations are reinforcing this trend. In complex product production systems (PPS), engineering changes are associated with extraordinary consequences. Engineering changes are generally considered to cause costs. In a study from 2012 with participants from industrial practice, 82% of the study participants indicated an increase in costs for change implementation [2]. One reason for this can be seen in the processing effort required, since technical changes take up an average of 25% of the total Research & Development (R&D) capacity [2]. Furthermore, with increasing complexity of the product, an increasing amount of subsequent changes can be expected [3].

Therefore, cost estimation of late introduced changed requirements for body shop systems has to be made. A decision is then made as to whether the change is to be absorbed by adapting the production line or whether design alternatives are to be examined [1]. Making appropriate decisions for defining and changing requirements, requires a holistic view. The scope of necessarily considered interdisciplinary requirements in the PPS extends along the entire PDP. This explicitly includes a production wise analysis, contrary to the classic product-sided understanding of requirements. The integration of production-side RM into existing, largely product-oriented processes offers opportunities to record and manage requirements in a structured manner and to incorporate them at the appropriate point in the development process as part of Concurrent Engineering [4]. This emphasizes the influence of a temporal perspective within RM, making a consideration before and after the definition of requirements necessary. Requirements must be examined a priori with regard to their consequence and evaluated a posteriori, in order to point out the consequences of selected solution mechanisms and at the same time make a priori investigations more reliable than established manual assessments [5]. Sustainable handling of requirements and engineering changes offers better planning capability with regard to time and equipment. The aim of this paper is to answer the research question:

How can existing approaches of RM and ECM in the literature be analyzed with respect to their usability in agile PDP?

2. Research Methodology

To answer the research question, an appropriate research framework is first established. The research procedure is based on the Design Research Methodology (DRM) of *Blessing et al.* [6]. Research clarification is intended to provide an overview and generate thematic understanding. Tasks of the research clarification consist in the definition of an overall research goal and in the identification of indications that support the developed assumptions. Relevant disciplines and the topic areas to be investigated need to be identified. Objectives include determining preliminary success criteria and identifying influencing factors that contribute to conceptualizing an adaptive engineering change system which enables agile PDP. A systematic literature review is then necessary to find approaches that represent the preliminary success criteria. The Research Clarification will serve as the basis for a later Descriptive Study. Factors that show the greatest influence are to be addressed for the subsequent Prescriptive Study. The results are the development of initial reference and impact models that describe the existing and the desired situation.

3. State of Research

The first step is to clarify the terms used in context with requirements and changes. In literature and in general use of language, a partly synonymous use of requirements and changes exists. Requirements and changes are therefore defined and explained in this context.

3.1 Requirements vs. Engineering Changes in PDP

A requirement is defined by the German Institute for Standardization (DIN) as "a need or expectation that is specified, usually assumed, or mandatory" [7]. The Institute of Electrical and Electronics Engineers (IEEE) characterizes a requirement as (1) a condition or capability needed by a user to solve a problem or achieve an objective. (2) A condition or capability that must be met or possessed by a system or system component to satisfy a contract, standard, specification, or other formally imposed documents. (3) A documented representation of a condition or capability as in (1) or (2) [8]. Further definitions for requirements can be found in [9]. For the formulation of requirements, rules exist, which deal with the use of suitable units, the avoidance of inaccurate designations, compound sentences as well as abbreviations, and the definition as well as consistency of acronyms [10]. Starting with change, it is generally defined by DIN as "the agreed specification of a new state in place of the previous state" [11]. *Jarrat et al.* point out that “change is defined as an alteration made to parts, drawings or software that have already been released during the product design process and life cycle, regardless of the scale or the type of the change” [12]. *Wickel* describes engineering changes as "modifications to products, product models, as well as related components, using a defined change process, after they have been released for further development and production" [13]. The existence of a release date distinguishes an engineering change from an iteration in the development process. In practice, the implementation of engineering changes may be linked to mandatory approvals [14]. Further definitions for engineering changes can be found in [3].

Considering the highlighted terminological relationship of the requirements with engineering changes, an engineering change implies to be a reaction on non-fulfilment of requirements and/or new definition of requirements [13]. There are procedures for handling of both requirements and engineering changes using RM and ECM. In agile PDP, however, the differently named domains show strong similarities, since every change is due to new requirements. A consideration of both domains therefore seems inevitable for the development of an adaptive engineering change system.

3.2 Handling of Requirements and Changes

The tools used for RM and ECM include proprietary software, generic reference processes and industry-specific standard processes. Proprietary software used for RM offers comprehensive options for managing requirement documents. There are possibilities to track changes and the effects of changes can be examined on the basis of previously generated links [9]. From a process perspective, the handling of changes can be retraced using generic processes. With a generic high-level ECM process, *Jarrat et al.* provide an often-cited basis for ECM approaches. Triggered by a change trigger, the reference process runs through the phases before, during, and after approval of the change in six steps. The elicitation of the change request is followed by an identification of possible solutions. This is followed by an assessment of the risks and impacts of solutions. After a solution is selected and approved by a change committee, a solution is implemented. The process is completed by a review [12]. An ECM reference process has been developed by the Automotive Industry Association (VDA) in collaboration with the Strategic Automotive Data Standards Industry Group to provide guidelines for the automotive industry in dealing with engineering change. The reference process includes the phases Identification of Potential for Change, Development of Alternative Solutions, Specification of and Decision for Change, Engineering Implementation of Change and Manufacturing, and Implementation of Change [15]. In addition to the aforementioned software tools and reference processes, industry- and company-specific standard processes are also common. For the automotive industry, a standard change process has proven its worth, helping in particular with the coordination of approval and reconciling processes. Adapted processes can also be found for other industries [16].

Deficiencies of such solutions are shown by a pronounced latency. Between the definition of the requirement, its assessment, and its ultimate implementation in production, there are further time-consuming process steps which cause this latency. This represents an inefficiency and obstacle in agile PDP. Further deficits can be seen in the mostly manual linking by a requirement engineer employed to coordinate the requirements. Due to the complex PPS resulting primarily from the large number of interdependent technologies, processes and logistic chains [16], the data sets to be examined and its complexity can become enormous [17], which severely limits the manageability for manual procedures. Additional complexity arises from the different times of requirement definition and from the interdisciplinary PDP, which makes it necessary to model the interdependencies. The amount of knowledge required for this exceeds the possibilities of manual evaluation quickly.

4. Sustainable Handling Through Data-based Methods

The development towards automated, data-based methods, which rely on decision making based on data, can enable sustainable handling of requirements and engineering changes. Individual components or even the entire RM or ECM process could be implemented more effectively using data-based methods. Data mining is one of the methods for combining data into information and converting it into knowledge. „Data mining is defined as the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful in that they lead to some advantage, usually an economic one. The data is invariably present in substantial quantities" [18]. The application of data mining techniques in ECM is currently still quite little researched [5]. However, the data required for this is certainly available, since quality management standards used in the automotive industry, such as DIN EN ISO 9001, require the retention of documented information [19]. Nevertheless, according to expert statements, the data are not used sufficiently [13].

In this context, the application of data mining or other artificial intelligence techniques offers possibilities to make predictions already during the creation of a change request. Another example are automatically generated suggestions on how to reduce the lead time of the change [20]. As a consequence, there can be the ability to provide a prediction of the impact of changes as instant feedback for the stakeholders involved in the PDP, thus simultaneously reducing the latencies present in usual ECM processes. A strong data basing requires sufficient modeling of the structures involved *(cf. chap. 3.2)*. Again, the high complexity of automotive PPS places correspondingly high demands on modeling. Changes are inevitable in the development of a complex product [21]. In order to effectively manage changes, it is of utmost importance to understand the effects, probability and propagation paths of engineering changes and thus gain capabilities for sustainable decisions [21]. An established variant of modeling products and product development processes is based on the use of characteristics and properties. Characteristics define the product, while properties describe the behavior of the product and result from the characteristics [22].

In terms of the described DRM, an initial impact model as a deliverable of the previous explanations is to be built up. **Figure 1** refers to the "Industrie 4.0 Maturity Index" in accordance to *Schuh et al.* [23], which identifies three stages on the way to an adaptive engineering change system starting from a computerized and connected PPS. The impact model is based on the levels of visibility, transparency, and predictive capacity, as computerization and connectivity are assumed to be the starting situation.

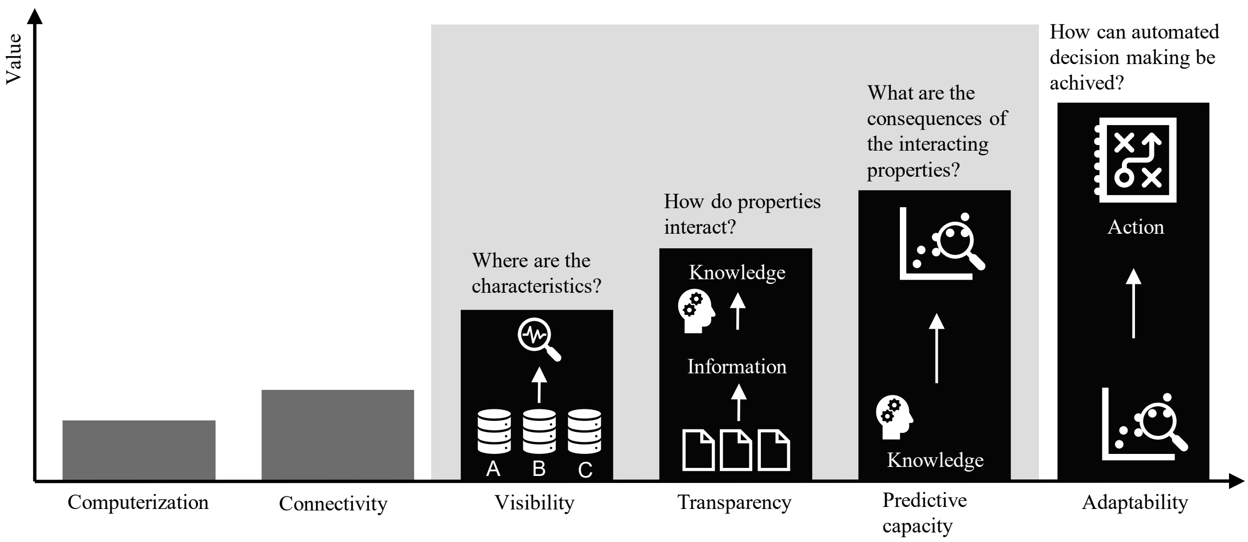


Figure 1: Integration into the Industrie 4.0 Maturity Index. According to *Schuh et al.* [23]

Within the stages, the levels of the Knowledge Pyramid can be classified. The Knowledge Pyramid places characters, data, information, knowledge, and action in a hierarchical relationship [24]. Features of the PPS must first be found in the data stored in decentralized silos. With regard to the interdisciplinary PDP, a comprehensive access must be created, and the visibility limited to individual departments must be expanded. The collected data is the basis for the step toward transparency, in which data is aggregated into information through semantic linking. This provides the opportunity to uncover interactions and thus interpret characteristics in terms of their properties, which determine the requirements through knowledge. This knowledge provides the basis for mapping the consequences of the properties and interactions. Future scenarios can be simulated through the acquired predictive capacity. Predictive capacity, in turn, is the basis for automated action and decision-making.

The model puts the topics visibility, transparency, and predictive capacity for building an adaptive engineering change system into focus. Thereupon, preliminary success criteria are now to be determined. To handle the required amounts of data, approaches characterized by a high degree of automation and *data basing* are required. To be able to justifiably make the implications of development decisions in the interdisciplinary PDP, approaches are required that take a holistic view of the PPS, including *product and production side*. In order to uncover properties and interactions in the PPS, modeling of structures is required. The *modeling of product structure and production structure* therefore represents a further criterion in the analysis of the approaches. A *consideration of interdependencies* within the PPS is necessary to link the structures on an abstracted level and to understand mutual influences. Both *a priori and a posteriori analysis* is required to ensure predictive capacity and to achieve learning effects. To consider the use case described in *chap. 1*, an approach that has already been tested for *body shop applications* is preferable.

5. Systematic Literature Review

Based on the six preliminary success criteria developed above, a systematic literature review of RM and ECM approaches was conducted.

5.1 Methodology For Literature Review

The STARLITE methodology was used to conduct the systematic literature review [25]. The STARLITE methodology provides guidelines for documenting literature searches in terms of sampling strategy, type of studies, approaches, range of years, limits, inclusions and exclusions, terms used, and electronic sources. A purposive, representative search strategy was chosen as the sampling strategy. The type of studies was limited to dissertations, papers, monographs, and articles in collective works. In addition to the structured literature search, an explorative approach based on the snowball principle was used for the initial search. The range of years was limited to entries from 2005 and until 2020. The search was limited to German and English language. Inclusions and exclusions were made by limiting the search to the technical field or mechanical engineering, as well as a target-oriented, individual adjustment of the filter categories and a focus of the search on title, keywords and abstract. In selecting the terms used, search components were determined and divided into main and secondary components. While the main components describe the thematic framework of RM and ECM, the secondary components address desired elements of the application domain, formalization, and methods. For the search, the databases Web of Science and Scopus were selected as electronic sources. In addition, the German National Bibliography, as well as SpringerLink and Hanser eLibrary were used. A search string was developed by composing the main and secondary components with AND and OR links. An example search string for Web of Science is shown below: *TS=(("Engineering Change Management" OR "Requirements Engineering“ OR “Product Changes”) AND (data* OR "Machine Learning" OR AI OR Body OR Automotive OR Manufacturing OR Production OR Development OR “Product Features” OR “Product Properties” OR “Product Characteristics” OR dependency)).* For the three last mentioned databases with limited search string functionality, a simple keyword search was used. After performing the initial search run, a large set of search results was obtained. The PRISMA Methodology was used to reduce this initial result to a manageable level [26]. After removing duplicates, the search results were reduced to 1.062 entries. 193 titles remain after selecting by title relevance. After a review of abstracts and conclusions, 71 entries were determined to be suitable for next step. These entries were further reduced after a full review, leaving 38 entries for final evaluation. An overview of the search results remaining for evaluation is shown in **figure 2.** The evaluation of the approaches is described in the following chapter.

	Criteria						Evaluation scheme		
							Consideration of product and production side: ●: both sides, ◐: production or product, ○: other **Consideration of the body shop:** ●: body shop, ◐: universal / automotive, ○: other **Consideration of PPS-Interdependencies:** ●: syst. representation, ◐: loose relations, ○: no representation **Structure modeling:** ●: syst. modeling, ◐: show possibilities, ○: no modeling **Degree of data basing:** ●: automated data-based, ◐: structured, ○: experience-based **A priori and a posteriori analysis:** ●: a priori and a posteriori, ◐: a priori or a posteriori, ○: no analysis		
Author	Consideration of product and production side	Consideration of the body shop	Consideration of PPS-Interdependencies	Structure modeling	Degree of data basing	A priori and a posteriori analysis	Level	Represents	Based on
Elezi 2011 [27]	●	◐	◐	◐	◐	◐	Framework	Data-based process design	KDD
Sharafi 2013 [28]	◐	◐	●	◐	●	●	Framework	Data-based process design	KDD
Jarrat 2011 [12]	●	◐	●	○	◐	●	Framework	ECM/RM process design / modeling	Engineering Change
Harms 2009 [14]	◐	◐	◐	◐	◐	◐	Framework	ECM/RM process design / modeling	Engineering Change
Wu 2014 [29]	◐	◐	○	○	◐	○	Framework	ECM/RM process design / modeling	Engineering Change
Kröger 2019 [30]	◐	◐	●	◐	○	◐	Framework	ECM/RM process design / modeling	Engineering Change
Brandl 2019 [31]	◐	◐	○	○	○	○	Framework	ECM/RM process design / modeling	Engineering Change
Basse 2019 [32]	●	●	●	◐	◐	◐	Framework	ECM/RM process design / modeling	Requirements
Meis 2017 [4]	●	◐	◐	◐	◐	●	Framework	ECM/RM process design / modeling	Requirements
Deubel 2007 [33]	●	◐	◐	●	◐	◐	Framework	Structure modeling	Characteristics
Weber 2001 [22]	●	◐	◐	●	◐	◐	Framework	Structure modeling	Characteristics
Köhler 2009 [3]	●	◐	●	●	◐	◐	Framework	Structure modeling	Characteristics
Grieco 2017 [34]	○	◐	◐	○	●	◐	Method	Clustering and searching	Automated
Kocar 2010 [35]	◐	○	●	●	●	◐	Method	Clustering and searching	Automated
Haunstetter 2010 [36]	◐	●	◐	○	●	○	Method	Clustering and searching	Automated
Wickel 2016 [13]	◐	◐	●	●	●	●	Method	Clustering and searching	Automated + Matrix
Schedl 2008 [9]	●	◐	◐	◐	◐	◐	Method	Documentation	Proprietary software
Ma 2017 [37]	◐	◐	●	◐	◐	◐	Method	Impact prediction	Graph-based
Koch 2016 [38]	◐	◐	◐	○	○	○	Method	Impact prediction	Graph-based
Wilms 2019 [39]	◐	◐	●	◐	○	◐	Method	Impact prediction	Graph-based
Plehn 2017 [40]	◐	◐	●	●	○	◐	Method	Impact prediction	Graph-based
Masmoudi 2020 [41]	◐	◐	◐	◐	○	◐	Method	Impact prediction	Graph-based
Hamraz 2013 [42]	◐	◐	●	●	◐	◐	Method	Impact prediction	Matrix-based
Hein 2018 [43]	◐	◐	●	◐	●	◐	Method	Impact prediction	Matrix-based
Leistner 2019 [44]	●	●	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Bauer 2015 [45]	◐	○	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Gärtner 2011 [46]	◐	◐	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Tang 2018 [47]	●	◐	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Kattner 2019 [48]	○	◐	◐	◐	◐	◐	Method	Impact prediction	Matrix-based
Giffin 2009 [17]	◐	◐	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Ullah 2018 [49]	◐	◐	●	◐	◐	◐	Method	Impact prediction	Matrix-based
Heimes 2020 [50]	◐	◐	◐	◐	◐	◐	Method	Impact prediction	Matrix-based
Lee 2017 [51]	◐	◐	●	◐	◐	◐	Method	Impact prediction	Statistical method
Bracht 2015 [52]	◐	◐	◐	○	◐	◐	Method	Impact prediction	Statistical method
Schweitzer 2020 [20]	◐	○	○	◐	◐	◐	Method	Impact prediction	Statistical + automated
Petzelt 2009 [53]	●	◐	◐	◐	◐	○	Method	Linkage	Feature-based
Bossmann 2007 [54]	◐	◐	○	●	◐	○	Method	Linkage	Feature-based
Hirani 2005 [55]	◐	◐	○	◐	◐	○	Method	Linkage	Graph-based

Figure 2: Evaluation of the identified approaches based on the systematic literature review

5.2 Description of the Evaluation Criteria

Based on the previously defined six criteria *(cf. chap. 4)*, a three-stage evaluation scheme was introduced that describes the respective degree of fulfillment of the criteria. Based on their fulfillment of the evaluation criteria, the approaches were classified into an evaluation matrix *(cf. fig. 2)*. The degree of fulfillment is represented through the three stages by Harvey Balls (full-half-empty). A full Harvey Ball corresponds to a good match with the respective evaluation criteria. The explanation of the degrees of fulfillment of the evaluation categories is shown in a legend in the figure.

5.3 Categorization

As shown in **figure 3**, the identified approaches address different levels in terms of their solution detailing. Some approaches represent comprehensive frameworks, each addressing different previously defined criteria. While differing in terms of their focus, the objectives of frameworks generally are to classify and relate. Overarching frameworks for ECM / RM process design and modeling are available, as are frameworks for data-based process design and approaches to modeling structures. Other approaches specifically apply certain methods that deal with clustering and searching, impact prediction and linkage and thus represent different parts within identified reference processes and the developed model. A method can be described as procedure to reach goals and is aiming to a lower level of abstraction. The methods presented are to be seen as enablers for the stages visibility, transparency and predictive capacity, while the frameworks form a basis for procedures. Therefore, to establish an adaptive engineering change system, the development of a holistic framework is required, which is designed with methods from different domains.

In the collection and clustering of requirements and engineering changes, as well as searching for best practice solutions, automated data-based approaches have been applied by using technologies such as Natural Language Processing and Sequential Pattern Mining. In predicting impact and propagation of changes, matrix-based and statistical approaches have been used in particular. Of particular note is the Design Structure Matrices-based Change Prediction Method by *Clarkson et al.* [56], which has been used and extended by many authors. Graph-based and feature-based approaches were found suitable for representing linkages in simple products or PPS. In addition, there are previously described software solutions and complementary research tools that combine some of the aforementioned methods. Many of the approaches shown are not initially designed for automated use, but nevertheless provide a basis for building data-based procedures on them.

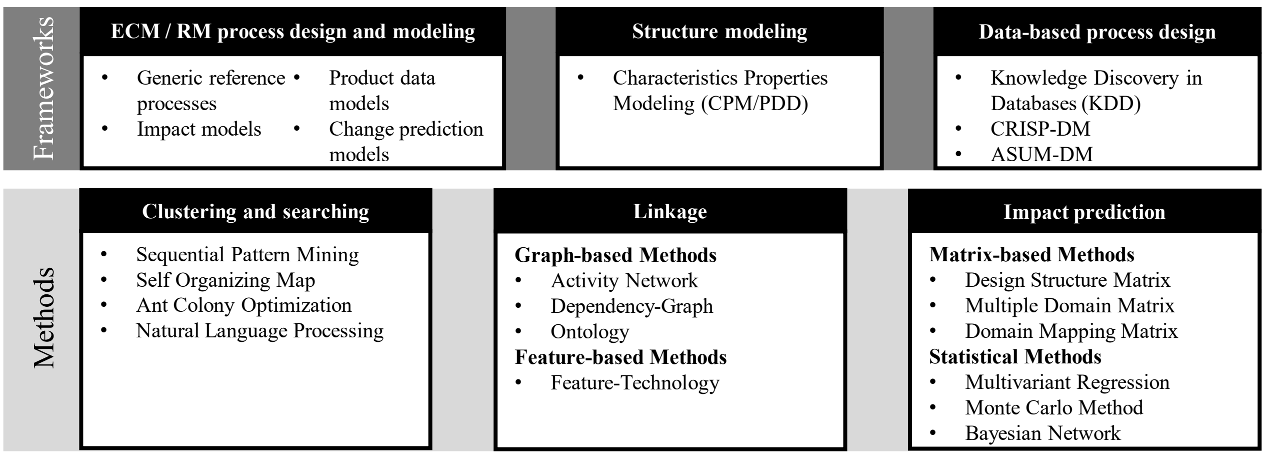


Figure 3: Categorization of identified frameworks and methods

6. Conclusion and Outlook

Parts of RM and ECM processes, that are gaining importance in agile PDP, can be designed more efficiently using data-based methods. Considering quality management standards, the data required for this should be available but is not used sufficiently due to a lack of formalization and modeling. An initial impact model towards sustainable handling of requirements and engineering changes was developed by showing a path to gain predictive capacity from data. Preliminary success criteria for an adaptive engineering change system were determined. An evaluation of approaches with respect to their usability in the present environment of an interdisciplinary, dynamic PDP was performed. Referring to the research question, approaches could be categorized based on the literature review. This analysis revealed the interdisciplinary nature of the use case as different domains have influence on modeling of the PPS. The results obtained offer solutions within their specific disciplines. However, for the applicability in agile PDP, an overlaying framework concept is still missing, which connects the identified methods and embeds the individual frameworks by forming an additional level above them. The methods serve as enablers for the stages visibility, transparency, and predictive capacity presented in the developed model towards an adaptive engineering change system. Predictive introduction of requirements based on data would make it possible to increasingly align the handling of engineering changes with the introduction of initial requirements and, at the same time, to reduce consequential changes. In the derivation of leading criteria, a large number of influencing factors must be taken into account. A variety of existing approaches provides the opportunity to form a holistic solution that represents a sufficient representation of applicable methods in an adequate framework. Therefore, it needs to be further developed by the authors in future research.

References

[1] Burr, H., 2008. Informationsmanagement an der Schnittstelle zwischen Entwicklung und Produktionsplanung im Karosserierohbau. Dissertation, Saarbrücken, 295 pp.

[2] Langer, S., Wilberg, J., Maier, A., Lindemann, U., 2012. Änderungsmanagement-Report 2012: Studienergebnisse zu Ursachen und Auswirkungen, aktuellen Praktiken, Herausforderungen und Strategien in Deutschland. Technische Universität München, 47 pp.

[3] Köhler, C., 2009. Technische Produktänderungen: Analyse und Beurteilung von Lösungsmöglichkeiten auf Basis einer Erweiterung des CPM/PDD-Ansatzes. Diss. Univ. des Saarlandes, Saarbrücken, 246 pp.

[4] Meis, J.-F., 2017. Produktionsseitiges Anforderungsmanagement. Diss., München, 220 pp.

[5] Tale-Yazdi, A., Kattner, N., Becerril, L., Lindemann, U., 2018. A Literature Review on Approaches for the Retrospective Utilisation of Data in Engineering Change Management, in: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, pp. 612–616.

[6] Blessing, L.T., Chakrabarti, A., 2009. DRM, a Design Research Methodology. Springer London, London.

[7] Deutsches Institut für Normung, 2015. Qualitätsmanagementsysteme – Grundlagen und Begriffe (ISO 9000:2015). Beuth, Berlin, 104 pp.

[8] The Institute of Electrical and Electronics Engineers, 1990. IEEE standard glossary of software engineering terminology, New York, 84 pp.

[9] Schedl, S., 2008. Integration von Anforderungsmanagement in den mechatronischen Entwicklungsprozess, München, 176 pp.

[10] International Council on Systems Engineering (INCOSE), 2012. Guide for Writing Requirements, San Diego, 59 pp.

[11] Deutsches Institut für Normung. Dokumentationssystematik (DIN 6789): Änderungen von Dokumenten und Gegenständen Allgemeine Anforderungen. Beuth, Berlin, 5 pp.

[12] Jarratt, T.A.W., Eckert, C.M., Caldwell, N.H.M., Clarkson, P.J., 2011. Engineering change: an overview and perspective on the literature. Res Eng Design 22 (2), 103–124.

[13] Wickel, M.C., 2016. Änderungen besser managen - Eine datenbasierte Methodik zur Analyse technischer Änderungen. Diss., München, 185 pp.

[14] Harms, E., 2009. Änderungs- und Konfigurationsmanagement unter Berücksichtigung von Verwendungsinstanzen: Arbeitsmethoden für integrierte Produktmodelle im Rahmen des Produkt-Lebenszyklus-Managements der Automobilindustrie. Diss., Karlsruhe, 250 pp.

[15] Verband der Automobilindustrie, 2009. Engineering Change Management Reference Process: White Paper. Strategic Automotive Product Data Standards Industry Group, 28 pp.

[16] Schuh, G., Stölzle, W., Straube, F., 2008. Anlaufmanagement in der Automobilindustrie erfolgreich umsetzen. Springer Berlin Heidelberg, Berlin, Heidelberg, 275 pp.

[17] Giffin, M., Weck, O. de, Bounova, G., Keller, R., Eckert, C., Clarkson, P.J., 2009. Change Propagation Analysis in Complex Technical Systems. Journal of Mechanical Design (131).

[18] 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.

[19] Deutsches Institut für Normung, 2015. Qualitätsmanagementsysteme – Anforderungen (ISO 9001:2015): Deutsche und Englische Fassung EN ISO 9001:2015. Beuth, Berlin, 71 pp.

[20] Schweitzer, G.M., Bitzer, M., Vielhaber, M., 2020. Produktentwicklung: KI-ready? ZWF 115 (12), 873–876.

[21] Ullah, I., Tang, D., Yin, L., 2016. Engineering Product and Process Design Changes: A Literature Overview. Procedia CIRP 56, 25–33.

[22] Weber, C., Werner, H., 2001. Schlussfolgerungen für „Design for X“ (DfX) aus der Perspektive eines neuen Ansatzes zur Modellierung von Produkten und Produktentwicklungsprozessen, Saarbrücken, 12 pp.

[23] Schuh, G., Anderl, R., Gausemeier, J., Hompel, M. ten, Wahlster, W. (Eds.), 2017. Industrie 4.0 Maturity Index: Managing the digital transformation of companies. Herbert Utz Verlag GmbH, München, 53 Seiten.

[24] Rowley, J., 2007. The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science 33 (2), 163–180.

[25] Booth, A., 2006. ‘‘Brimful of STARLITE’’: toward standards for reporting literature searches. Journal of the Medical Library Association JMLA 94 (4), 421–430.

[26] Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ (Clinical research ed.) 339, b2535.

[27] Elezi, F., Sharafi, A., Mirson, A., Wolf, P., Krcmar, H., Lindemann, U., 2011. A Knowledge Discovery in Databases (KDD) Approach for Extracting Causes of Iterations in Engineering Change Orders, in: Volume 2: 31st Computers and Information in Engineering Conference, Washington, DC, USA. ASMEDC, pp. 1401–1410.

[28] Sharafi, A., 2013. Knowledge Discovery in Databases. Springer Fachmedien Wiesbaden, Wiesbaden, 326 pp.

[29] Kröger, S., Bauer, H., Reinhart, G., 2019. Kostenstruktur zur Bewertung von Änderungsauswirkungen in der Produktion. ZWF 114 (9), 530–534.

[30] Wu, W.-H., Fang, L.-C., Wang, W.-Y., Yu, M.-C., Kao, H.-Y., 2014. An advanced CMII-based engineering change management framework: the integration of PLM and ERP perspectives. International Journal of Production Research 52 (20), 6092–6109.

[31] Basse, F., 2019. Gestaltung eines adaptiven Änderungssystems für einen beherrschten Serienhochlauf. Diss., Aachen, 255 pp.

[32] Brandl, F., Stahlmann, J.-T., Gunther Reinhart, 2019. Digitalisierung des Änderungsmanagements: In drei Schritten zur automatisierten Änderungsprüfung. ZWF 114 (11), 766–770.

[33] Deubel, T., 2007. Anforderungs-, kosten- und wertgetriebene Steuerung des Produktentwicklungsprozesses. Diss. LKT Univ, Saarbrücken, 166 pp.

[34] Grieco, A., Pacella, M., Blaco, M., 2017. On the Application of Text Clustering in Engineering Change Process. Procedia CIRP 62, 187–192.

[35] Haunstetter, T., 2010. Methoden des Data Mining in Anwendung an die Flexibilitätsattribute des automobilen Karosseriebaus zur strategischen Investitionsplanung. Diss., Hamburg, 173 pp.

[36] Kocar, V., Akgunduz, A., 2010. ADVICE: A virtual environment for Engineering Change Management. Computers in Industry 61 (1), 15–28.

[37] MA, S., JIANG, Z., LIU, W., HUANG, C., 2017. Design Property Network-Based Change Propagation Prediction Approach for Mechanical Product Development. Chin. J. Mech. Eng. 30 (3), 676–688.

[38] Koch, J., Gritsch, A., Reinhart, G., 2016. Process design for the management of changes in manufacturing: Toward a Manufacturing Change Management process. CIRP Journal of Manufacturing Science and Technology 14, 10–19.

[39] Wilms, R., Cemmasson, V.F., Inkermann, D., Reik, M., Vietor, T., 2019. Identifying Cross-Domain Linkage Types to Support Engineering Change Management and Requirements Engineering. Procedia CIRP 84, 719–724.

[40] Plehn, C., 2017. A method for analyzing the impact of changes and their propagation in manufacturing systems. Dissertation.

[41] Masmoudi, M., Zolghadri, M., Leclaire, P., 2020. A posteriori identification of dependencies between continuous variables for the engineering change management. Res Eng Design 31 (3), 257–274.

[42] Hamraz, B., Caldwell, N.H., Wynn, D.C., Clarkson, P.J., 2013. Requirements-based development of an improved engineering change management method. Journal of Engineering Design 24 (11), 765–793.

[43] Hein, P.H., Voris, N., Morkos, B., 2018. Predicting requirement change propagation through investigation of physical and functional domains. Res Eng Design 29 (2), 309–328.

[44] Leistner, B., Angelov, D., Mayer, R., 2019. Generic model for production-related requirements and their interdependencies on the example of automotive chassis. Automot. Engine Technol. 4 (3-4), 179–188.

[45] Bauer, W., Nepomuk, C., Lindemann, U., Maurer, M., 2015. Domain-Spanning Change Propagation in Changing Technical Systems, in: Cardin, M.-A., Krob, D., Lui, P.C., Tan, Y.H., Wood, K. (Eds.), Complex Systems Design & Management Asia. Springer International Publishing, Cham, pp. 111–123.

[46] Gärtner, T., 2011. Simulationsmodell für das Projekt- und Änderungsmanagement in der Automobilentwicklung auf Basis der Design Structure Matrix. Diss. Shaker, Aachen, 280 pp.

[47] Tang, D., Yin, L., Ullah, I., 2018. Matrix-based Product Design and Change Management. Springer Singapore, Singapore.

[48] Kattner, N., Mehlstaeubl, J., Becerril, L., Lindemann, U., 2018. Data Analysis in Engineering Change Management – Improving Collaboration by Assessing Organizational Dependencies Based on Past Engineering Change Information, in: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, pp. 617–621.

[49] Ullah, I., Tang, D., Wang, Q., Yin, L., Hussain, I., 2018. Managing engineering change requirements during the product development process. Concurrent Engineering 26 (2), 171–186.

[50] Heimes, H.H., Kampker, A., Haunreiter, A., Davids, H., Klohs, D., 2020. Product-requirement-model to approach the identification of uncertainties in battery systems development. Int J Interact Des Manuf 14 (3), 911–922.

[51] Lee, J., Hong, Y.S., 2017. Bayesian network approach to change propagation analysis. Res Eng Design 28 (4), 437–455.

[52] Bracht, U., Geckler, D., 2015. Stabile Projekte durch verbessertes Änderungsmanagement.

[53] Petzelt, D., Schallow, J., Deuse, J., Ferstl, H., 2009. Produktionsgerechte Produkte durch technische Mitgestaltung aus der Produktionsplanung. ZWF 104 (11), 988–992.

[54] Bossmann, M., 2007. Feature-basierte Produkt- und Prozessmodelle in der integrierten Produktentstehung. Diss. Univ. des Saarlandes Lehrstuhl für Fertigungstechnik/CAM, Saarbrücken, 190 pp.

[55] Hirani, H., Ratchev, S., 2005. Knowledge-Based Requirements Engineering for Reconfigurable Precision Assembly Systems, in: Camarinha-Matos, L.M. (Ed.), Emerging Solutions for Future Manufacturing Systems, vol. 159. Kluwer Academic Publishers, Boston, pp. 359–366.

[56] Clarkson, P.J., Simons, C., Eckert, C., 2004. Predicting Change Propagation in Complex Design. Journal of Mechanical Design 126 (5), 788–797.

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 in Aachen and director of the Research Institute for Rationalization (FIR) at the RWTH Aachen. He is the founder of Schuh & Co. Group and the e.GO Mobile AG. Prof. Schuh is a member of several supervisory boards and boards of directors.

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 research associate at the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

Christian Ahues (*1996) is a student at the department of Vehicle Production at Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

The Acceptance Of The Blockchain Technology In Food Supply Chains – A Literature Review

Dino Hardjosuwito[1], Themo Voswinckel[1], Eva Bausch[1]
[1]Institute for Industrial Management (FIR) at RWTH Aachen, Aachen, Germany

Abstract

The blockchain technology has been increasingly applied in industrial use-cases in recent years. Although the food industry fits in particular with the requirements for blockchain applications, since the actors barely know each other and trust plays a crucial role, it is not widely established in the food industry. There are efforts to increase transparency and enable traceability in food supply chains by applying blockchain technology to share data in a trustworthy way across companies and to ensure food quality standards. This technology can be further used to enable the identification of inconsistencies in sensor data and more efficient handling of food recalls across the food supply chain. The success of a new technology depends to a large extent on its acceptance by companies and their employees. This paper deals with the acceptance of such a blockchain application and presents a systematic literature review to summarize the methods and results of acceptance analyses of the blockchain technology in food supply chain s. Particular attention has been devoted to traceability. For this objective, research is analysed based on scientific methods and the results are systematically analysed.

Keywords

Blockchain; Acceptance analysis; Food supply chain; Literature review

1. Introduction

Blockchain technology is increasingly being applied in a variety of industries. Many different participants in a supply chain lead to distinctive transactions between them. These transactions can be reflected very well by a blockchain.[1] Blockchain technology offers several important benefits especially for the supply chain, such as traceability, efficiency, and transparency.[2]

Especially in food supply chains, the previously mentioned properties have a major role to play. Here, blockchain technology can be used to comply with food standards. Thus, it can help to achieve more control over the heterogeneous, complex and dynamic food supply chains. This control is now also desired by more and more consumers in order to obtain more safety and quality. [3] Furthermore, there are some regulatory and social constraints to provide a transparent supply chain. This can avoid possible scandals that have occurred in the food industry in the recent past and create more trust between the different parts of the food supply chain and especially the customer. [4]

Despite these clear benefits, blockchain technology has almost no practical application. In fact, many projects have been discontinued. [3] Similarly, in a survey by Statista, only 36% of decision-makers surveyed said that they had so far investigated the topic area of blockchain technology. [5]

Most research on blockchain technology to date has been related to technical implementation. However, it is also important that there is a certain level of acceptance so that the technology that has been developed is also used. [2]

Therefore, the purpose of this paper is to summarize the methods and results of the acceptance analysis of blockchain technology in the food supply chain to this point. It is important that projects and companies involved in blockchain technology know which models and factors they need to consider in order for the technology to be accepted from the market and companies. Otherwise, more and more projects will not continue. However, if they know what models to use, they can analyse their own market well and can respond to the particular needs of the market to create more acceptance. In addition, this paper already provides the first factors that can have a positive or negative impact on certain markets.

For this purpose, a systematic literature review is conducted below. In the following the exact procedure of the literature research is described. Subsequently, the results are categorized, and a part is further analysed to finally come to a summary and discussion of the results.

2. Research Model

To identify the methods and results of an acceptance analysis of blockchain technology in the food supply chain, a systematic literature review has been conducted. The detailed process of this search is presented in the following. Several databases were searched to obtain a more comprehensive search scope. It has been decided to choose the databases "Web of Science"[1] and "Scopus"[2]**.** This database was chosen in order to obtain the most scientifically sound results possible. In addition, special setting options are included in the mentioned databases.

2.1 Implementation

For "Web of Science" the following search settings have been defined. First, it has been determined that all databases are searched to achieve the most comprehensive result as possible. Secondly, it has been specified that the total available time period is to be searched, this refers to the period from 1926-2021. No restrictions have been made for the language, whereby there were only English language search results because of the English keywords. In addition, the search setting "topic" was selected, in which the papers were sorted according to the specified search criteria in the title, abstract, author keywords, and more were scanned according to the given search criteria. Furthermore, only open accessible literature was included in the research.

Similar settings were used for "Scopus". There were no settings for database selection or language, whereby there were also here only English-language search results. In addition, the available total time period for “Scopus” was considered again, what includes the period from 1960-2021. The type of database search was defined for “Scopus” as well. The settings "Article title, Abstract, Keywords" were chosen to receive a search result as accurately as possible and to have a similar search algorithm as in "Web of Science".

Subsequently, the databases were searched for specific search criteria. At "Scopus" no further adjustments of the search criteria are audible. At " Web of Science" however, there is the opportunity to put words that stand together in quotation marks, these are then seen as a total phrase. However, to have the same search conditions in both databases, this setting was not used in this paper. The found papers are then sorted by citations, so that scientifically relevant papers are always analysed as part of the search.

---

[1] www.webofknowledge.com
[2] www.scopus.com

After all formal configurations have been completed, the first search term "Acceptance Blockchain food supply chain" has been specified. This is the first search term, as it suits the problem best. The term "acceptance" refers to the acceptance analysis as defined in the problem, "blockchain" defines the utilization of the blockchain technology and "food supply chain" describes the area of utilization.

Since the previously described search is a very specific application area, a second search was subsequently conducted. This search is no longer concerned with the specific application area of the food supply chain, but only looks at the supply chain as a general area. Therefore, the new search term will be " Acceptance blockchain supply chain".

2.2 Findings

With the first mentioned search term, five papers could be found each at "Web of Science" and "Scopus". After removing the duplicates, six papers remained. Whereby it can be assumed that with other databases it would not have been possible to find many more papers, since there were already so many duplicates in the two databases which were considered.

After a closer look at the abstract and the conclusion, there were only two papers left that matched the given research problem. The others discussed the blockchain technology, but not necessarily in a food supply chain and were not concerned with an acceptance analysis at all. Since this is a too small number for a comprehensive literature review, the search has been repeated with a new defined search term. In order to be able to make a well-founded analysis of the papers and to scientifically evaluate the applied methods as well as the results, the search was repeated with a newly defined search term.

By screening the databases with the second search term, 43 papers have been found at "Web of Science" and 46 papers have been found at "Scopus". This time, only papers that have already been cited at least once have been analysed, so that only papers that have already been scientifically recognized are analysed. Therefore, 28 papers from Scopus and 35 papers from Web of Science have been analysed.

Again, the abstracts and conclusions were evaluated to determine the relevance of the paper for this literature review. The same problem occurred as in the previous search, that not all papers address an acceptance analysis. Some of them do mention that possible acceptance needs to be considered, but do not address this issue further. Due to this problem, after a closer analysis, only 14 papers are left that are worth to be analysed further in the next chapter.

Table 1 shows an overview of the papers found during the search. The papers from the first search are greyed out. Here, the following content criteria have been considered: "Blockchain Technology", "Acceptance Analysis", "Supply Chain". This means that it must be mentioned in the abstract or in the conclusion that the paper deals with the given topics. So that in the next chapter only the relevant papers are analysed in more detail. All three topics must be covered in one paper in order for it to be evaluated further.

Table 1: Results of both Searches

Author	Year	Paper	Blockchain Technology	Acceptance	Supply chain	Used	Database	Citations
Abu-elezz et. al	2020	The benefits and threats of blockchain technology in healthcare: A scoping review	x				Scopus, Web of Science	10
Alazab et. al	2021	Blockchain technology in supply chain management: an empirical study of the factors affecting user adoption/acceptance	x	x	x	x	Scopus, Web of Science	5
Alkhodre et al.	2019	A Blockchain-based value added tax (VAT) system: Saudi Arabia as a use-case	x		x		Scopus, Web of Science	7
Amin& Zuhairi	2021	Crowdfunding Smart Contract: Security And Challenges	x				Scopus	0
Astill et. al	2019	Transparency in food supply chains: A review of enabling technology solutions			x		Scopus, Web of Science	39
Busse et. al	2019	A response to the united nations cites blockchain challenge: Incremental and integrative poa-based permit exchange	x		x		Scopus	1

Caldarelli& Ellul	2021	Trusted Academic Transcripts on the Blockchain: A Systematic Literature Review	x				Scopus	1
Choi et. al	2020	Factors Affecting Organizations' Resistance to the Adoption of Blockchain Technology in Supply Networks	x	x	x	x	Scopus, Web of Science	3
Choo et. al	2020	Blockchain Ecosystem—Technological and Management Opportunities and Challenges	x				Scopus, Web of Science	2
Clohessy et. al	2020	Antecedents of blockchain adoption: An integrative framework	x	x	x	x	Web of Science	2
Durach et. al	2020	Blockchain Applications in Supply Chain Transactions	x	x	x	x	Web of Science	11
Falcone et. al	2020	Understanding Managers' Reactions to Blockchain Technologies in the Supply Chain: The Reliable and Unbiased Software Agent	x	x	x	x	Web of Science	1
Ferri et. al	2020	Ascertaining auditors' intentions to use blockchain technology: evidence from the Big 4 accountancy firms in Italy	x	x			Web of Science	3
Gökalp et. al	2019	Acceptance of Blockchain Based Supply Chain Management System: Research Model Proposal	x	x	x	x	Scopus	1
Gupta et. al	2020	Prioritizing intentions behind investment in cryptocurrency: a fuzzy analytical framework		x			Web of Science	2
Jain et. al	2020	Blockchain in logistics industry: in fizz customer trust or not	x	x	x	x	Scopus, Web of Science	2
Jovic et. al	2020	Improving Maritime Transport Sustainability Using Blockchain-Based Information Exchange	x	x			Web of Science	5
Kamble et. al	2021	A machine learning based approach for predicting blockchain adoption in supply Chain	x	x	x		Scopus, Web of Science	2
Kamble et. al	2019	Understanding the Blockchain technology adoption in supply chains-Indian context	x	x	x	x	Scopus, Web of Science	129
Karamchandani et. al	2020	Perception-based model for analyzing the impact of enterprise blockchain adoption on SCM in the Indian service industry	x		x		Scopus, Web of Science	16
Kouhizadeh et. al	2021	Blockchain Technology and the Sustainable Supply Chain: Theoretically Exploring Adoption Barriers	x	x	x	x	Scopus, Web of Science	36
Min	2019	Blockchain technology for enhancing Check for updates supply chain resilience	x		x		Web of Science	95
Nawaz& Thowfeek	2020	Blockchain technology adoption by chain professionals	x	x	x		Scopus	2
Park	2020	A Study on Sustainable Usage Intention of Blockchain in the Big Data Era: Logistics and Supply Chain Management Companies	x	x	x	x	Scopus, Web of Science	1
Queiroz et. al	2020	Blockchain adoption in operations and supply chain management: empirical evidence from an emerging economy	x	x	x	x	Scopus, Web of Science	13
Queiroz& Wamba	2019	Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA	x	x	x	x	Scopus, Web of Science	173
Rainero& Modarelli	2021	Food tracking and blockchain-induced knowledge: a corporate social responsibility tool for sustainable decision-making	x		x		Scopus, Web of Science	0
Sander et. al	2018	The acceptance of blockchain technology in meat traceability and transparency	x	x	x	x	Web of Science	40
Schuetz& Venkatesh	2020	Blockchain, adoption, and financial inclusion in India: Research opportunities	x	x	x		Web of Science	34
Srivastava	2019	Blockchain and transaction processing time using M/M/1 queue model	x				Scopus	1
Supranee& Rotchana-kitumnuai	2017	The acceptance of the application of blockchain technology in the supply chain process of the Thai automotive industry	x		x		Scopus	10
Tan& Sundarakani	2021	Assessing Blockchain Technology application for freight booking business: a case study from Technology Acceptance Model perspective	x	x			Scopus, Web of Science	2
Thiruchelvam et. al	2018	Blockchain-based technology in the coffee supply chain trade: Case of Burundi coffee	x	x	x		Scopus	15
Tribia et. al	2018	Supply chain management based on blockchain: A systematic mapping study	x		x		Scopus	22

Upadhyay	2020	Demystifying blockchain: A critical analysis of challenges, applications, and opportunities	x		x		Web of Science	11
Wamba et. al	2020	Dynamics between blockchain adoption determinants and supply chain performance: An empirical investigation	x		x		Web of Science	18
Wamba& Queiroz	2020	Industry 4.0 and the supply chain digitalisation: a blockchain diffusion perspective	x		x		Web of Science	8
Wamba& Queiroz	2019	The role of social influence in blockchain adoption: The Brazilian supply chain case	x	x	x	x	Scopus, Web of Science	3
Wang et. al	2019	Designing a Blockchain Enabled Suplly Chain	x		x		Web of Science	3
Wang et. al	2019	Making sense of blockchain technology: How will it transform supply chains?	x		x		Web of Science	96
Wong et. al	2020	Unearthing the determinants of Blockchain adoption in supply chain management	x	x	x	x	Web of Science	18
Wong et. al	2020	Time to seize the digital evolution: Adoption of blockchain in operations and supply chain management among Malaysian SMEs	x	x	x	x	Scopus, Web of Science	53
Wu et. al	2021	An analysis of strategies for adopting blockchain technology in the fresh product supply chain	x		x		Scopus, Web of Science	1
Yang	2019	Maritime shipping digitalization: Blockchain-based technology applications, future improvements, and intention to use	x	x			Scopus, Web of Science	45
Zhou et. al	2020	From traceability to provenance of agricultural products through blockchain	x		x		Scopus, Web of Science	0

3. Results

After the initial classification of the papers into relevant and non-relevant, the relevant papers were analysed in more detail in the following. Particular attention was paid to the methods used in the individual papers and the results they arrived at. The models applied and the corresponding factors are explained in more detail. In order to achieve this, the papers are first briefly summarized before the results are discussed further.

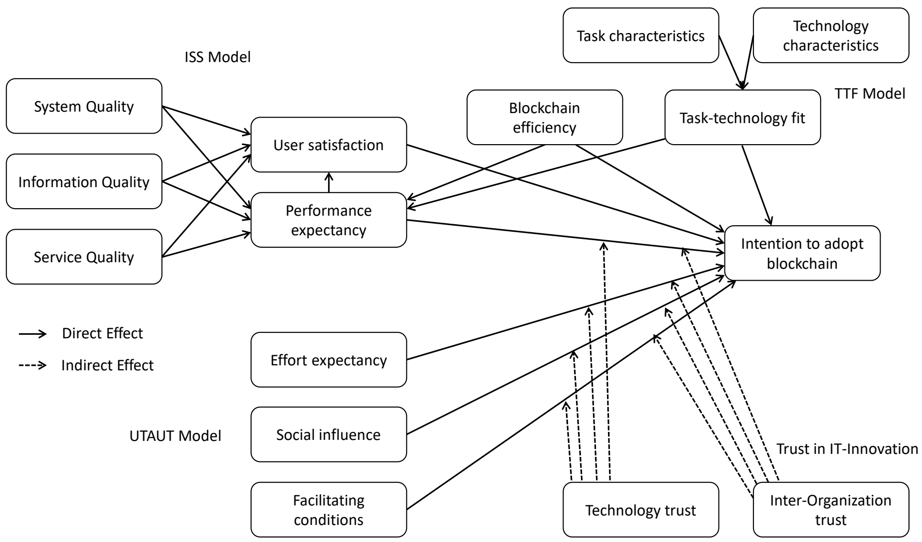
The first paper that was considered in more detail is "The acceptance of Blockchain technology in meat traceability and transparency" by Sander at al. from 2018. This paper will be further just referred to as Paper 1. In this paper, stakeholder acceptance in the meat industry is investigated under the positive influence of traceability and transparency. Hypotheses have been formulated first. To verify the previously formulated hypotheses, 712 questionnaires have been sent by mail, with 141 responses. In addition, 12 semi-structured interviews were conducted with supply chain stakeholders. The study concluded that consumers are overwhelmed by the present situation, as there exist too many different certificates on the market, the true meaning of a certificate is unclear. This is where blockchain technology can significantly improve the situation. In addition, it is also assumed that traceability will increase the quality of food. However, some government representatives have concerns about the financing and feasibility of the system. [6]

The 2nd paper, "Blockchain Technology and the Sustainable Supply Chain: Theoretically Exploring Adoption Barriers," by Kouhizadeh in 2020, deals with the potential issues that may occur during the adoption of blockchain technology. For this purpose, the Force Field Theory is considered together with the TOE model. TOE is a theory that describes how 3 factors influence the acceptance of a new technology with the T represents Technology, the O represents Organization, and the E represents Environment. The selected sample includes 47 interviewees. They have been given a questionnaire, which has been previously prepared with experts. The created questionnaire aims to compare the factors: "Organizational Barriers, Internal Supply Chain Organizational Barriers, Technological Barriers, External Barriers" in pairs. The result of the questionnaire indicated that the Technological Barriers and the Internal Supply Chain Organizational Barriers were evaluated the highest. However, it can also be concluded that there is a correlation between the two barriers. Therefore, it can be assumed that the Technological Barriers must be solved before the

Internal Supply Chain Organizational Barriers. In addition, it can be observed that the External Barriers also need to be solved before the other barriers in order to successfully adopt the blockchain technology. [2]

The next paper of the search is "Blockchain technology in supply chain management: an empirical study of factors affecting user adoption/acceptance" by Alazab et al. from 2020, this paper is further named Paper3. This paper focuses on the most important influences of the acceptance analysis. For this analysis, hypotheses have been formulated regarding the factors of Figure 1. To investigate these hypotheses, a questionnaire was created. For this purpose, the hypotheses have been categorized in the unified theory of acceptance and use of technology (UTAUT), information system success (ISS) and task- technology fit (TTF) method. Thus, the model shown in Figure 1 is created. The questionnaires were then sent to 2000 managers from 184 companies in Australia. The authors received 449 responses from 104 different companies. After analysing the results, the authors concluded that employee perspectives or other experiences of other users do not influence the decision of utilizing blockchain technology. Instead, what matters is interorganizational and technical trust. Additionally, trust between different actors in the supply chain also matters a lot. [7]

Figure 1: Model with UTAUT, ISS and TTF [7]



In the next section, we will look at the paper "Factors Affecting Organizations Resistance to Adoption of Blockchain Technology in Supply Networks" by Choi at al. from 2020, which will be referred to as Paper4 later in this paper. This paper considers the barriers to blockchain technology adoption based on the TOE model. Hypotheses are designed using the TOE model, which are to be verified. To create a questionnaire, additional information from different papers and publications of well-known companies are also included. The Hypotheses are designed utilizing the TAM model. The TAM model, like the previously mentioned models, describes factors that affect a decision to adopt a new technology. The designed questionnaires are disseminated via social media so that 92 questionnaires are completed. These are then analysed, and the authors concluded that technical maturity, cost, compatibility, and scalability are the biggest barriers to the successful adoption of the blockchain solution. There are also security and privacy concerns. However, the reduced time required to collaborate with other companies is perceived positively. In addition, a high level of expertise has a positive influence on the acceptance of blockchain technology. [8]

The Paper 5 "A study on sustainable usage intention of blockchain in the Big Data Era: Logistic and Supply Chain Management Companies" by Park from the year 2020 focuses on the acceptance of blockchain technology in the supply chain based on the UTAUT and TOE model. The UTAUT model considers gender, age, experience, and voluntariness of use in addition to performance expectancy, effort expectancy, social influence and facilitating conditions, although the first factors are not considered further in the paper. The TOE model has already been explained in more detail in the previous summaries. Based on these models, several hypotheses were formulated. These have been subsequently investigated by reference to a literature review and a questionnaire survey. The survey included 800 questionnaires, which were distributed, 172 were returned completed. After analysing the results, it was found that the factors of the UTAUT model as well as the TOE model show a significant positive impact. [1]

The 2019 paper "Time to seize the digital evolution: adoption of blockchain in operations and supply chain management among Malaysian SME's" by Wong et al. is the 6th paper to be analysed in more detail. This paper is about the adoption of blockchain technology among SMEs in Malaysia. In this paper, hypotheses have been formulated again based on the TOE model. This time, the hypotheses focus not only on the behaviour of the companies but also on how the Technical Effects influence the Organizational Effects. For this purpose, a questionnaire is again created, whereby 194 completed questionnaires can be evaluated. The authors have concluded that the cost savings in working with third parties have a positive effect on the decision to use blockchain technology. In the same way, the total acquisition cost is seen as an attraction. On the other hand, the complexity of the system is a negative factor, as it leads to fear of utilization. Competitive pressure is also seen as a positive factor, as its use can lead to a competitive advantage. However, the support of the management level is still missing. [9]

In the next section the 2020 paper "Understanding Managers' Reaction to Blockchain Technologies in the Supply Chain: The Reliable and Unbiased Software Agent" by Falcone et al. is reviewed, this paper is further named Paper 7. In this paper, the acceptance of software agents in the supply chain is considered. Here to different areas are examined: Cognitive stage, affective stage, behaviour stage. The cognitive stage deals with blockchain design features, the affective stage with perceptions, which in turn are divided into risk, trustworthiness, and justice, and the behaviour stage with the willingness to use. On this basis, hypotheses were formulated, which were tested with 141 fully completed questionnaires. The authors concluded that there is an overall positive attitude toward thist technology. Here, the factors of the affective stage promote the willingness to use blockchain technology. In addition, the new technology is seen as embedded, reliable, and biased. [10]

The paper "Blockchain in logistics industry: in fizz customer trust or not" by Jain et. al. from 2019, attempts to explain the behaviour of an individual regarding to blockchain technology, this paper is further named Paper 8. The TAM model has been utilized as the basis of hypothesises formulated. Here, the following factors have been considered in more detail: Attitude, Perceived Ease of Use, Perceived Usefulness. Based on the previously formulated hypotheses, a questionnaire was created and distributed to 250 online shoppers. The authors returned 240 completed questionnaires. From this, they were able to conclude that the previously defined factors all have a significant influence. In addition, they expect higher business revenues and better customer relations. [11]

Consider the 2020 paper "Unearthing the determinants of Blockchain adoption in supply chain management" by Wong et. al, this paper is further named Paper 9. This paper focused on the decision making regarding blockchain technology by Malaysian companies. For this purpose, hypotheses have been formulated based on the UTAUT model. However, the UTAUT model has been extended by the factors Technology Readiness and Technology Affinity. A questionnaire was then prepared. This questionnaire was then answered by 162 Malaysian companies. The authors concluded that Performance Expectancy and Effort Expectancy are insignificant for the decision. In contrast, the Facilitating condition is significant, as well as the two additional factors. However, the respondents saw difficulties with regulations and guidelines, especially in data protection and consumer protection. However, the respondents also have little experience in blockchain technology. Therefore, the authors believe that greater awareness and expertise of the technology, as well as clear regulations, are helpful and necessary. [12]

The 2020 published paper "Blockchain adoption in operations and supply chain management empirical evidence from an emerging economy“ by Queiroz et. al, focuses on the decision-making process of Brazilian companies to use blockchain technology, this paper is further named Paper 10. For this purpose, questionnaires were prepared based on the previously mentioned UTAUT model. These were then sent via social media. The authors received 184 responses. From these replies, the authors concluded that facility conditions, trust, and effort expectancy have a positive impact on decision making, and performance expectancy has a negative impact. [13]

Furthermore, the paper "Blockchain-Based Supply Chain Management: Understanding the Determinants of Adoption in the Context of Organization" by Gökalp from 2020 was also examined in more detail, this paper is further named Paper 11. For this purpose, certain factors were clustered with the help of the TOE model. The individual clusters were as follows: Technology: relative advantage, complexity, compatibility, interoperability, standardization, scalability, and trust; Organization: organizations IT resource, top management support, organization size, financial resources; Environmental: competitive pressure, trading partner pressure, government policy, government regulations, inter-organizational trust. Interviews were conducted with 30 experts on these factors to obtain a weighting of the factors. The environment is the most important factor for the exchange of information between organizations. The technical factors are particularly concerned with improving performance and efficiency, and the organizational factors are concerned with the availability of IT resources and finances. It is also about having the support of top management and inter-organizational trust, as well as supportive policy regulations and competitive advantage. [14]

The 2019 paper "The Role of Social Influence in Blockchain Adoption: The Brazilian Supply Chain Case" by Wamba & Queiroz analyses Behavioural Intention Adoption based on the UTAUT Models. This paper will be further just referred to as Paper 12. For this purpose, hypotheses on the factors Social Influence, Facility Conditions, Performance Expectancy, and Effort Expectancy have been formulated first. Based on these hypotheses, questionnaires were created and sent to Brazilian senior supply chain specialists via LinkedIn. The authors received 138 responses. They concluded that Social Influence has the strongest influence. However, Performance Expectancy has no influence, which they explained by the fact that managers in India do not expect any increase in productivity from blockchain technology. [15]

The next paper "Special Topic Forum: Blockchain Application and Strategies for Supply Chain Research and Practice" by Durach et. al is from 2021. This paper will be further just referred to as Paper 13. In this paper, the authors surveyed 151 German managers regarding their attitudes toward blockchain technology. Of particular importance here are the factors that are categorized under the keywords: Initial consideration, Active evaluation, Moment of purchase, Post purchase experience and Loyalty loop. They concluded that the cost of the technology plays a major role. The same applies to the handling and transfer of data. On the other hand, the factors loyalty, competitive pressure and transparency of processes have less impact. [5]

In addition, the paper "Understanding the Blockchain technology adoption in supply chains- Indian context" by Kamble & Gunasekaran from 2018 was analysed in more detail. This paper will be further just referred to as Paper 14. In this, the TAM, TPB (Theory of planned behaviour) and TRI (Technology readiness index) models are used to create hypotheses about the adoption of blockchain technology. These hypotheses serve as the basis for a survey with 150 participants from 150 companies and the design of a questionnaire, which was completed by 181 supply chain professionals. The authors concluded that there are no concerns about implementation or ease of use. In addition, the perceived benefits are also positive. There is a discernible influence of the subjective norm which can manifest itself in the form of peer pressure or a competitive advantage. In addition, personal attitude also has an influence on decisions. [16]

In addition, the paper "Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA" by Queiroz & Wamba from 2018 is also considered. This paper will be further just referred to as Paper 15. This compares the adoption of blockchain technology in the US and India. For this purpose, hypotheses are formulated based on the TAM and UTAUT models. These serve as the basis for questionnaires completed by 394 supply chain professionals in the US and 344 in India. They show that Performance Expectancy always has an influence. Trust, on the other hand, has a greater influence in India than in the USA. Whereas Facilitating conditions has a greater influence in the US. Social Influence, however, is again stronger in India, where the influence of colleagues or family on decisions is greater.[17]

The last paper, "Antecedents of blockchain adoption: An integrative framework" by Clohessy et al. from 2020, this paper is further named Paper 16, looks at decision making to utilize blockchain technology considering the TOE model. However, this was extended to include the Task and Individual categories. After a literature review, the authors concluded that the categories considered are relevant to decision making. As well as the support of top management. [18]

Taking a closer look at the analysed papers, it is noticeable that mainly 3 models are applied, the TOE model, TAM model and UTAUT model. Table 2 summarizes these models and the corresponding factors.

Table 2: Models and Factors

Model	Factors	Meaning
TOE	Technological factors	Availability, Characteristics
	Organisational actor	Formal and informal Linking Structures, Communication Processes, Size, Slack
	Envoirmental factors	Industry Characteristics and Market Structure, Technology Support Infrastructure, Goverment Regulations
TAM	Perceived usefulness	"the degree to which a person believes that using a particular system would enhance his or her job performance"
	Percieved ease of use	"degree of belief that a person will find it easy to use the system"
	Attitude	"an individual's positive or negative feelings (evaluative affect) about performing the target behavior"
UTAUT	Performance expectancy	"the degree to which an individual believes that using the system will help him or her to attain gains in job performance"
	Effort expactancy	"the degree of ease associated with the use of the system"
	Social influence	"the degree to which an individual perceives that important others believe he or she should use the new system"
	Facility conditions	"the degree to which an individual believes that an organizational and technical infrastructure exist to support the use of the system"
	Gender	control variables
	Age	
	Experience	
	Voluntariness	

The TOE model has three main categories under which factors are collected. In the table, some factors are mentioned, but there are no fixed factors that can be added. The TAM model is limited to the final factors that influence the decision. There are also extensions like the TAM2. In this model the factors that influence the performance expectancy, Effort expactency and Attitude are also listed. The UTAUT model only considers the factors listed in the table, these are not further classified. In this model, however, there are still control variables, which are not always used.

All models have different emphases, which is why it can be useful to use them in combination. Paper 5 and Paper 15 e.g., have considered the TAM and the UTAUT model together. Other authors like Paper 3 and Paper 14 have extended the models by further not so frequently used models. This approach is often useful in order to consider a further spectrum of factors.

4. Conclusion

During this literature review, it was noticed that there are not many papers on the acceptance analysis of blockchain technology in the food supply chain yet. For the acceptance of blockchain technology in the supply chain, there are already a few more papers, but even here there is still a lot of research to be done. The most noticeable thing is that the research is restricted to the last few years. The oldest paper that we have analysed is from 2018 but most are from 2020. This is mainly because the technology used is new, however, this field of research should not be underestimated. However, it has been noted that new papers have been published on this topic during our research period.

It can be concluded that the acceptance of the blockchain technology depends on a wide range of factors. The possible traceability and transparency are seen as positive. In addition, blockchain technology increases the trust of the various actors in the supply chain and reduces the effort and costs incurred by third parties. Furthermore, social factors and competitiveness play a major positive influence. The age and role of the decision-makers also have a positive impact on the decision. Negative factors are the financing of blockchain technology. In addition, the complexity and compatibility are viewed critically. In addition, despite the high technological standard, security and especially data security are considered. However, one must always pay attention to which market one is in, so it was found that the acceptance factors can differ with regard to the country in which research was conducted. [17] These Factors are also summarised in Table 3.

Table 3: Factors for the utilization of blockchain technology

Paper	Factors for the utilization of blockchain technology	Factors against the utilization of blockchain technology
Paper 1	traceability will increase the quality of food, more transparent than the current situation	financing, feasibility of the system
Paper 2		internal supply chain organizational barriers, technological barriers
Paper 3	Interorganizational and technical trust, trust between different actors	
Paper 4	time required to collaborate with other companies, high level of expertise	technical maturity, cost, compatibility, and scalability
Paper 5	UTAUT model and the TOE model have s significant positive impact	
Paper 6	working with third parties have a positive effect, acquisition cost, competitive pressure	complexity of the system, fear of utilization, support of the management level is still missing
Paper 7	embedded, reliable, and biased	risk, trustworthiness, justice
Paper 8	higher business revenues and better customer relations	
Paper 9	facilitating condition, technology readiness and technology affinity	regulations and guidelines
Paper 10	facility conditions, trust, and effort expectancy	
Paper 11	performance, efficiency, environment. IT resources, finances, top management, inter-organizational trust, policy regulations and competitive advantage	
Paper 12	social influence	
Paper 13	pay for performance, transfer of contracts, bonds, deeds, or stocks, escrow services document-signing processes, token-curated registries	
Paper 14	subject norm, perceived usefulness, attitute	
Paper 15	performance expactancy, trust, social influence, facility conditions	
Paper 16	technik, organisation, environment	

The objective of this paper was to summarize different methods and results of acceptance analysis of blockchain technology in food supply chains. Although there have not been many papers on blockchain technology in food supply chains, we believe that the methods and factors found can be applied to food supply chains to find out whether the technology is accepted or not. And thus, more successful projects can be implemented in this field.

References

[1] Park, K.O., 2020. A Study on Sustainable Usage Intention of Blockchain in the Big Data Era: Logistics and Supply Chain Management Companies.

[2] Kouhizadeh, M., 2020. Blockchain Technology and the Sustainable Supply Chain: Theoretically Exploring Adoption Barriers. International Journal of Production Economics.

[3] Behnke, K., Janssen, M., 2020. Boundary conditions for traceability in food supply chains using blockchain technology. International Journal of Information Management 52, 101969.

[4] Voswinckel, T., 2020. Impact Analysis of Industrial Standards on Blockchainsfor Food Supply Chain.

[5] Durach, C.F., Blesik, T., Düring, M., Bick, M., 2021. Blockchain Applications in Supply Chain Transactions. J Bus Logist 42 (1), 7–24.

[6] Sander, F., 2018. The acceptance of blockchain technology in meat traceability and transparency.

[7] Alazab, M., 2020. Blockchain technology in supply chain management: an empirical study of the factors affecting user adoption/acceptance.

[8] Choi, D., 2020. Factors Affecting Organizations' Resistance to the Adoption of Blockchain Technology in Supply Networks.

[9] Wong, L.-W., 2019. Time to seize the digital evolution: Adoption of blockchain in operations and supply chain management among Malaysian SMEs.

[10] Falcone, E.C., 2020. Understanding Managers' Reactions to Blockchain Technologies in the Supply Chain: The Reliable and Unbiased Software Agent.

[11] Jain, G., Singh, H., Chaturvedi, K.R., Rakesh, S., 2020. Blockchain in logistics industry: in fizz customer trust or not. JEIM 33 (3), 541–558.

[12] Wong, L.-W., Tan, G.W.-H., Lee, V.-H., Ooi, K.-B., Sohal, A., 2020. Unearthing the determinants of Blockchain adoption in supply chain management. International Journal of Production Research 58 (7), 2100–2123.

[13] Queiroz, M.M., Fosso Wamba, S., Bourmont, M. de, Telles, R., 2020. Blockchain adoption in operations and supply chain management: empirical evidence from an emerging economy. International Journal of Production Research, 1–17.

[14] Gökalp, E., Gökalp, M.O., Çoban, S., 2020. Blockchain-Based Supply Chain Management: Understanding the Determinants of Adoption in the Context of Organizations. Information Systems Management, 1–22.

[15] Wamba, S.F., Queiroz, M.M., 2019. The Role of Social Influence in Blockchain Adoption: The Brazilian Supply Chain Case. IFAC-PapersOnLine 52 (13), 1715–1720.

[16] Kamble, S., Gunasekaran, A., Arha, H., 2019. Understanding the Blockchain technology adoption in supply chains-Indian context. International Journal of Production Research 57 (7), 2009–2033.

[17] Queiroz, M.M., Fosso Wamba, S., 2019. Blockchain adoption challenges in supply chain: An empirical investigation of the main drivers in India and the USA. International Journal of Information Management 46, 70–82.

[18] Clohessy, T., Treiblmaier, H., Acton, T., Rogers, N., 2020. Antecedents of blockchain adoption: An integrative framework. Strategic Change 29 (5), 501–515.

Biography

Dino Hardjosuwito (*1988) is a research associate at the Institute for Industrial Management (FIR) at the RWTH Aachen in the research group Production Planning. He studied mechanical engineering, specializing in chemical engineering, as well as production engineering at the RWTH Aachen University. Currently, he is working on the research project “The Safe Food Chain by using blockchain technology (SiLKe)”, which aims to enable transparency and traceability of food supply chains.

Themo Voswinckel (*1992) is a research associate at the Institute for Industrial Management (FIR) at the RWTH Aachen in the research group Production Planning and a head of the taskforce “Blockchain for industrial applications”. He studied business administration and engineering, specializing in materials and process engineering, at the RWTH Aachen University. Currently, he is working among others on the research project “Safe Food Chain by using blockchain technology (SiLKe)”, which aims to enable transparency and traceability of food supply chains.

Eva Bausch (*1996) is a research assistant at the Institute for Industrial Management (FIR) at the RWTH Aachen. She studies business administration at the RWTH Aachen University. Currently, she is working on the research project “The Safe Food Chain by using blockchain technology (SiLKe)”, which aims to enable transparency and traceability of food supply chains.

2nd Conference on Production Systems and Logistics

Dedicated Data Sovereignty as Enabler for Platform-Based Business Models

Sander Lass[1], Benedict Bender[1]
[1] Chair of Business Informatics, esp. Processes and Systems,
University of Potsdam, Germany

Abstract

The digitalization of value networks holds out the prospect of many advantages for the participating companies. Utilizing information platforms, cross-company data exchange enables increased efficiency of collaboration and offers space for new business models and services. In addition to the technological challenges, the fear of know-how leakage appears to be a significant roadblock that hinders the beneficial realization of new business models in digital ecosystems. This paper provides the necessary building blocks of digital participation and, in particular, classifies the issue of trust creation within it as a significant success factor. Based on these findings, it presents a solution concept that, by linking the identified building blocks, offers the individual actors of the digital value network the opportunity to retain sovereignty over their data and know-how and to use the potential of extensive networking. In particular, the presented concept takes into account the relevant dilemma, that every actor (e. g. the machine users) has to be able to control his communicated data at any time and have sufficient possibilities for intervention that, on the one hand, satisfy the need for protection of his knowledge and, on the other hand, do not excessively diminish the benefits of the system or the business. Taking up this perspective, this paper introduces dedicated data sovereignty and shows a possible implementation concept.

Keywords

information gateway; data security; information flow control; platform acceptance

1. Introduction

Digitalization provides technologies and concepts for realizing the integration of business-relevant information across the entire value chain. In particular, cross-company information platforms are a general approach in this respect. In conjunction with the Internet of Things (IoT), there is potential for manufacturing companies to ensure their competitiveness in the future. IoT-based systems and their integration by means of information platforms make it possible to link a large number of data sources and network them into advantageous value creation networks. Cross-company networking in production has already demonstrated its fundamental potential in the automotive industry [1]. By participating in platform ecosystems, companies can encounter cost pressure in the core business. Digital business models allow internal product and process optimization for the platform players machine suppliers, machine users, suppliers and the customer.

With numerous IoT-based platforms and various clouds, solutions exist that are already available as a usable product, at least from the perspective of their providers [2], [3]. Hyperscalers (AWS, Microsoft Azure and Google Cloud Platform) also offer basic functions in terms of computing power, storage and networks. Looking at actual use in practice, the potential users do not seem to share this optimism: deployment occurs hesitantly and use of information platforms, despite their potential, is low.

DOI: https://doi.org/10.15488/11299 publish-Ing.

The result is a low level of information exchange between value creation partners and level of digitization across entire industrys [4], [5]. Since companies as actors in value chains are shaping socio-technical systems, possible reasons are the perceived risks which, from the companies' point of view, have not yet been addressed to a sufficient extent and are working against actual comprehensive use. According to Gartner, IoT platforms are in the phase through of disillusionment [6]. The task is to identify solutions based on key challenges that will help to implement digitized value networks in a targeted manner. Taking up this perspective, this article introduces dedicated data sovereignty and shows a possible variant of implementation.

1.1. Research question and expected results

Experience with cloud-based ERP shows that the fear of a know-how leak is a relevant roadblock among potential users [7]. This leads to the thesis that trust creation plays a significant role and that ensuring perceptible data sovereignty is therefore an important success factor for the effective establishment of future value creation networks. The implication of this thesis is that the cross-company exchange of information as a fundamental part of digital value networks must be designed in such a way that the participating companies do not have to fear a leakage of their know-how. This makes controlled communication indispensable. The dilemma: It must be ensured that every actor (e.g., the machine users) can control the communicated data at any time and at an effort-adequate level, and that there are sufficient options for intervention that satisfy the need for protection on the one hand and do not excessively diminish the system benefits on the other. For SMEs in particular, this is a challenging task that has to be mastered in addition to day-to-day business.

Thus, there is a need for suitable concept of principles that serves as a basis for the development of such solutions. Solutions that avoid the perceived loss of control if data is passed on to external systems (e.g., cloud-based services or SaaS applications) and still make the aforementioned positive aspects usable. Solutions that create trust through the appropriate regulation of information flows bring acceptance for the active co-design of future ecosystems. In this respect, the following relevant research questions arise with regard to the establishment of platform-based business models in future value creation networks: **Is the fear of know-how outflow actually a misson-critical obstacle? How can the data sovereignty of individual actors within shared platforms be achieved with reasonable effort?** Subsidiary results from the knowledge process are: What is the significance of trust creation for the establishment of digital value networks? Which building blocks are success factors for connection to IoT platforms?

Taking up the above thesis, the identification of the obstacles confirms the high importance of trust creation within value networks. A reference model of the necessary building blocks systematizes the implementation of the platform connection. In this context, the reference model is to be understood as a model which, on the one hand, concretizes certain aspects of the mapping space (delimitation of metamodel) and, on the other hand, offers room for adaptation to concrete use cases (delimitation of model). It serves as a basis for a systematic approach and forms a framework for further operationalization to concrete use cases. With the goal of effective trust creation, we introduce dedicated data sovereignty, a concept that ensures appropriate know-how protection. Part of the dedicated data sovereignty is the concept of a controlling entity - an information firewall. This element acts as part of the corporate network and provides extensive control of information flow. This avoids the creation of another external actor that appears to be independent, but results in respect to its viable business model in additional cost and effort. The Information-Firewall provides the concept which takes over the tasks of information flow control within the corporate structure. Synergistically, the device also offers the prospect of further added value (e.g. retrofit).

1.2. Methodology of research process

The first part of the research process is the problem analysis with an identification and structuring of the obstacles that hinder the use of cross-company information platforms and platform-based business models. It starts with investigating the lack of participation in cross-company information platforms, carried out in the first instance by means of interviews with companies to identify practical obstacle reasons. For the purpose of consolidation and supplementation, a focused literature research follows in the second instance. The latter forms part of the state of Research. It follows by the elaboration of trust creation as an essential key factor follows. The resulting findings are then used to identify relevant fields of action. Based on this, the development of a solution approach follows. This begins with the data privacy approach and shows the current state of research in this area. Subsequently, security-relevant model elements are determined and linked in the concept of information firewalls. In addition, an implementation proposal results, which carries out the operationalization and enables the implementation for existing systems.

2. Problem analysis

The analysis begins with a practical perspective, considers possible solutions and derives implications. Thus, it links practical needs and theoretical solutions to a model of the relevant solution modules.

2.1. Practitioner's perspective

The situation from a practical perspective is documented in the results of 18 interviews with players in manufacturing SMEs (managing directors, IT managers, production managers; end of 2019). The guideline-based design allows sufficient freedom in the choice of areas of observation and their detailing, while still maintaining the focus. With regard to the use of platforms, the guideline covers the perception of the potentials, the experiences to date and the status of implementation, viewed from the perspectives of people, technology and organization. In summary, the following findings result.

The existing heterogeneous IT and automation landscape and the lack of integration capability of the production objects make the holistic implementation of platform projects difficult. There are too few standardized interfaces that allow the configuration and comprehensive connection of the systems involved in a way that is commensurate with the effort involved and overcomes the partial closeness of proprietary isolated solutions. Implementation also fails because of the specifics of the situation at hand. Although there is consensus on the theoretical application potential, difficulties arise in the individual adaptation. The selection and configuration of the appropriate technological and organizational elements and their sustainable combination represent hurdles. This also includes the lack of migration strategies that enable systematic and targeted further development of the existing systems (brown field), as well as the provision of human resources, since all employees are typically tied up in day-to-day business in SMEs.

Furthermore, the individual benefits of digitizing processes cannot be adequately demonstrated without suitable evaluation tools for potential investment decisions. The only partial evaluation of possible solution modules leads to false expectations and misjudgments on the part of those responsible and decision-makers. In addition to these technical and organizational reasons, the psychosocial dimension is also part of the problem with the fear of a know-how drain. Since control of the data actually communicated to the platform during operation cannot be adequately ensured, these concerns lead to the decision not to connect to and use the platform. In summary, significant obstacles exist within the enabling prerequisite, the implementation as well as the operation. This general structuring into this issue areas is applied in the following for further elaboration.

2.2. Detailing the issue areas

The next step of the knowledge process aims at a supplementary consideration of the reasons for obstacles by means of literature research and their systematization. The guiding question of the literature research is that of building blocks that make an essential contribution within the issue areas mentioned.

From a technological point of view, IoT systems arise from the connection of people, objects and systems [8], which interact as actors in a communication structure. Industry 4.0 as the production-related manifestation of the IoT names cyber-physical systems (CPS) and their extensive networking as essential elements that use embedded systems (ES) to equip objects with the necessary capabilities and upgrade them to IoT devices [9], [10]. Consequently, high penetration is a prerequisite for the economic implementation of CPS and their linkage to cyber-physical production systems (CPPS). The same applies to other technologies (e.g., AutoID, localization with GPS or beacons, algorithms for search and analysis such as deep learning) that realize essential capabilities of the IoT system elements. Thus, **technology availability** emerges as part of the issue area enabling prerequisite.

Networking to form an IoT communication architecture implies three potential levels of action for the system elements involved: environmental interaction with sensing by sensors and action by actuators, the gateway level as an essential network element, and the IoT platform as a higher level of data storage and processing [11]. Grounded in their ability to process information locally, IoT devices can not only be used for mere data collection, but can also act as an IoT gateway, if necessary, which handles communication to the next higher level (typically a cloud) (e.g., via HTTP/REST-based data transfer or via MQTT protocol) [12]. These gateways are equipped with various network interfaces and, in addition to pure data transmission, can also act as translators or intermediaries and perform preprocessing of the data (e.g., filtering, aggregation) locally [13]. Within the cloud, the cross-element evaluation of the accruing data and integration of the results into the company's business processes takes place. Possible variants are the private cloud on premise in the company or the external variant using cloud service providers [14]. Existing IT architectures must therefore have a suitable architecture concept that fundamentally permits this distribution of tasks for the system elements mentioned. **Architecture concepts** forms a further component of the issue area enabling prerequisite.

Taking up the above IoT communication structure, IoT platforms are software systems that connect objects or devices [15] and provide or allow the development of applications for data storage, analysis or visualization. To structure the architecture of IoT, typical models concretize the three task domains into a five-layer model with typically device, connectivity, processing, application, and security layers [16]. The latter three form the IoT platform in a narrow sense. The processing layer includes device management (with identification and health monitoring) and data preparation, among others. The application layer provides applications for information retrieval, some of which cover domain-specific use cases or can be created by external actors. Also part of an IoT platform is a security concept. Other reference models make further differentiation into functions and tasks or directly include business models and processes (e.g., ten task areas [14], eleven criteria [17]) and thus offer a broader perspective. In existing platform offerings, the aforementioned tasks are implemented or configurable to varying extents. Particularly relevant components are standardized interfaces to third-party systems, IoT data analytics and the provision of mobile applications, as well as support for a wide range of communication protocols [18]. This shows success-critical factors for the implementation of IoT platforms. The component IoT platform offer, i.e. **portfolio available on the market**, is also part of the enabling prerequisites issue area.

For the demonstration of use and validation of potential solutions, a suitable set of tools for evaluation is needed. The demand for such testing and validation tools is also evident from the increasing address in the relevant funding programs of science and practice as well as in political and industrial initiatives: As so-called "test centers", they are an essential part of the respective intended solution strategy [19], [20]). In terms of the structuring used, the **solution evaluation** is located in the implementation issue area.

In particular, the topics of interfaces and communication protocols point to the technical challenges of integration within heterogeneous and evolved IT and automation landscapes. Since it cannot be assumed that existing systems will be replaced by new copies without further ado, the implication arises that existing systems must be enabled in a suitable form to act as part of an IoT platform. In particular, the integration of closed legacy systems forms a typical use case [21], [22]. Reasons are the necessary investment protection, the new acquisition of machines or systems is rarely a real option for SMEs. An implementation within existing systems is inevitable. Thus, **brownfield** is part of the problem area implementation.

Various middleware concepts (such as the Reference Architecture Model Industry 4.0 (RAMI 4.0) [23]) exist to enable existing production facilities. These work on a very abstract level and offer little help for operational implementation. Thus, these general concepts require further concretization (e.g., in the form of configurable migration strategies). They must also address the specifics of the situation at hand, but at the same time follow standards in order to avoid isolated solutions or lock-in effects. Target-oriented **individualization** is another building block within the issue area of implementation.

The use of the Internet Protocol (IP) in the IoT gives the impression of simple implementation of global end-to-end communication. Looking at the differentiation into signal, data, information and aggregation levels [24], this is largely true only on the first two levels. On the information or aggregation level, this is not the case. The inclusion of semantic aspects or complex data structures, preprocessing and aggregation, or the provision of functions at a high level of abstraction for communication and operation is not sufficiently available and requires further concerted work [25]. This means production objects can only exchange data, but no full interoperability has been implemented at the application level of the IoT structural model in a practical way. Currently, there are only islands of interoperability provided by the individual reference architectures from different application domains [26]. In other words, objects just talk but do not understand each other sufficiently. After all, some commonly used application-level protocols have been established for connecting machines to cloud services, which provide a basis for solving the lack of communication capabilities. These include MQTT (Message Queue Telemetry Transport) for message exchange between devices, LWM2M (Lightweight Machine-to-Machine) for IoT device management, and OPC-UA (Open Platform Communication Unified Architecture) for communication between machines. Thus, **semantic interoperability** is also part of the implementation problem domain.

The commitment of the stakeholders is an important success factor in the introduction and use of innovations and counteracts possible negative consequences such as the deliberate delaying or slowing down of the change or even the failure of the project [27]. Acceptance is also characterized by the perception of risks and barriers when (potential) customers use IoT and cloud computing [28]. In particular, the assessment of the trustworthiness of the provider is a critical factor for success in this respect [29]. The fact that shop-floor IT is no longer an isolated entity that can only be accessed physically means that new threat scenarios are emerging (see Stuxnet, Duqu, etc.). As a consequence, new challenges arise for security concepts and their practical implementation in the factory [30]. On the other hand, a security concept must protect the know-how about production processes and manufacturing methods from uncontrolled outflow [31], which is also essential in the ecosystem [32]. Know-how protection is particularly important due to the high degree of networking of the IoT and its need for communication relevant to its use. Suitable measures are not only important in the initial design (cf. security by design). Rather, in daily operation, maintenance and necessary adaptation are highly relevant. For this reason, **know-how assurance** is classified as an essential component of subject area operation, as is **user acceptance** especially its upholding, which promotes a high degree of utilization of the system element in question.

The required ability to adapt to changing conditions during operation addresses adaptability. As the adaptation to changing conditions by the system itself, it is, in addition to efficiency, a further requirement with regard to the competitiveness of companies [33]. Concepts of adaptability provide more suitable design means and solution paths for permanent and rapid adaptation of the internal organization and technology [34]. Also in the context of digitalization, platform-based ecosystems and new business models, mutability is a desirable property of the overall enterprise system, whose architecture forms the basis for changing business models from within and changes due to competition and new technologies from outside [35]. Consequently, the system as well as the system element must offer internal user suitable options for implementing future change requirements in an effort-appropriate manner. Since not all change requirements are foreseeable during implementation, **adaptability** is taken into account as a further component of the issue area "operation". Figure 1 summarizes the resulting elements from the identified obstacle reasons and influencing factors, structured into the three issue areas prerequisites/enablers, implementation and operation. Also included are the perceived levels of maturity, which was part of the insight validation with the interview partners.

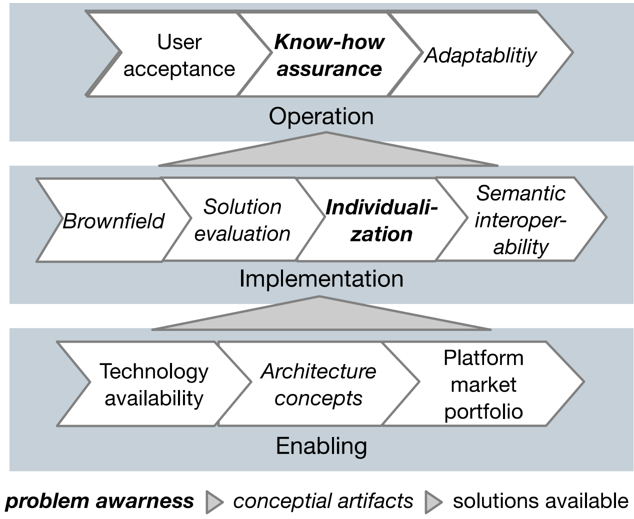


Figure 1: Building blocks for platform application and maturity

2.3. Implications

In line with the fact that psychosocial aspects are also relevant within change management for the actual use of innovations in addition to the technical and organizational perspective, concerns and worries on the part of potential users are important for acceptance and the probability of benefits. In the case of external information platforms, this relates in particular to fears of a loss of know-how. To counter these, it must be ensured that every actor (e.g., the machine users) can control the communicated data at any time and that there are sufficient intervention options which, on the one hand, satisfy the need for protection and, on the other hand, do not excessively diminish the system benefits. Existing authorization concepts must be checked in this respect or adapted to the extended data availability. A suitable data structuring with associated rights and appropriate data granularity is still missing.

The control of the actually communicated data is an essential consequence. The implication for action is therefore to ensure appropriate data sovereignty with respect to the platform connection by means of controlled communication that is transparent to the data owner at all times. Ensuring data sovereignty becomes an effectiveness-determining task in the use of information platforms and the realization of platform-based business models in this regard. Furthermore, this task does not have a "static" problem solution, but requires an adaptable solution that implements control loops, if necessary, in order to be able to make adjustments and further developments during operations and to react appropriately to new external and internal requirements. A particular challenge for SMEs is both the implementation and operation of these control loops, which implies the use of appropriate automation and configuration. These control loops need an effective runtime environment within the overall architecture and external knowledge of the business transactions at the meta-model level.

3. Dedicated data sovereignty

The particular challenge is to restrict the flow of information without impairing the functions of the platform ecosystem, i.e., without too much or too little data flowing out of the company. Thus, suitable mechanisms are required that on the one hand realize the fundamental data sovereignty, but on the other hand also address the information needs of the network partners in an expedient manner and adapt the communicated data to this purpose. This tailored data sovereignty is referred to as dedicated data sovereignty. The dedicated data sovereignty approach takes up the influencing factors that have been identified and derives a reference model of the necessary building blocks for implementation.

3.1. Starting point data privacy

Trust creation as a basis for acceptance and application is well recognized. Trust and Privacy is one of the key challenges in respect of the adoption of Internet of Things [36], wich is the collective term for future connecting system concepts. Likewise, the Gaia-X initiative of the BmWI points to trust, digital sovereignty and self-determination as relevant goals and recommends addressing them within modern cross-company data infrastructures [37]. In this respect, this initiative names digital sovereignty as a part of the implementation. An essential element here is data sovereignty, which concretizes the goal: complete control over stored and processed data and also the independent decision on who is permitted to have access to it.

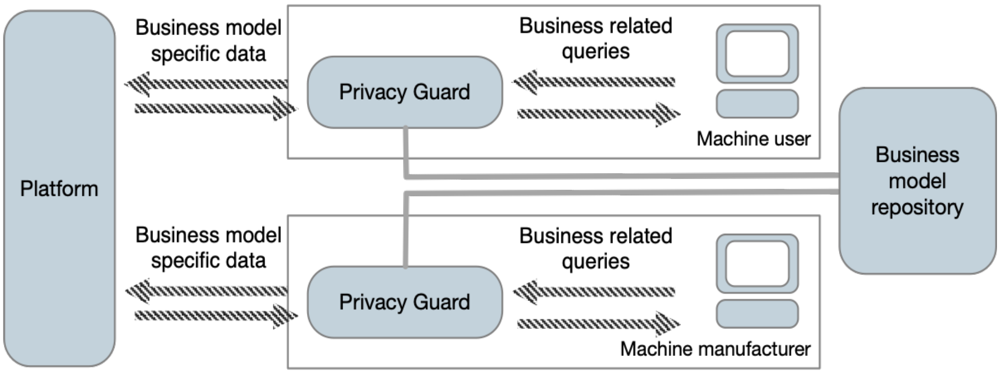


Figure 2: Implementation by using information gateways with a firewall function

Starting point data privacy: There are already approaches from data protection (e.g. BSI-Grundschutz) which can serve both methodically and in terms of content as a basis for developing an adequate adaptation to the IoT in the area of discourse being worked on. Various principles and paradigms exist here that serve as the starting point for further conceptualization of the reference model. A first operationalization is provided by [38]. The basic principle is data sparseness (minimize) with the restriction to the provision of the actually required data of the respective business activity. This implies the sufficient definition and delimitation of the respective use cases. In addition, encryption and anonymization (hide) ensure secure communication and appropriate information reduction. Distributed data storage and analysis (seprate) can reduce the risk of knowledge leakage by scattering information fragments, since the complete picture is not fully accessible to anyone, as can early aggregation into groups (aggregate) through local data aggregation. Both measures require appropriate data classification. Organizationally, transparency with regard to data collection, processing and dissemination as well as loss through attacks (inform) and the maintenance of control by the data owner (control) must be realized. The enforcement of data protection laws (enforce) and the demonstration of enforcement (demonstrate) also have an effect in this sense and may have a regulatory requirement. All building blocks are to be anchored in the architecture through technical and organizational measures (privacy by design).

The requirements of data economy in combination with the diversity of platform-based business transactions give rise to the need for scalable anonymization, i.e., the implementation of different degrees of anonymization and pseudonymization. Approaches are provided, for example, by Marnau with k-anonymity [39] and Ulbricht [40]. The differential privacy approach (cf. [41]) implies the multi-level design using gateway-like software elements. The gateway element allows security and pre-processing at the user's end. Figure 2 shows this approach. It extends the classic firewall with business related abilities, replacing static blocking of information flows with customized business case dependent filtering. The question remains open as to how the gateway element must be designed and how it can be ensured with regard to the requirement of appropriate integration and low-effort operation. Building on these basic approaches results in the concept of the information firewall. This concept of an information-firewall provides a solution to control the communicated data regarding actor-specific requirements for high transparency of the data flows as well as necessary interventions by the respective stakeholders.

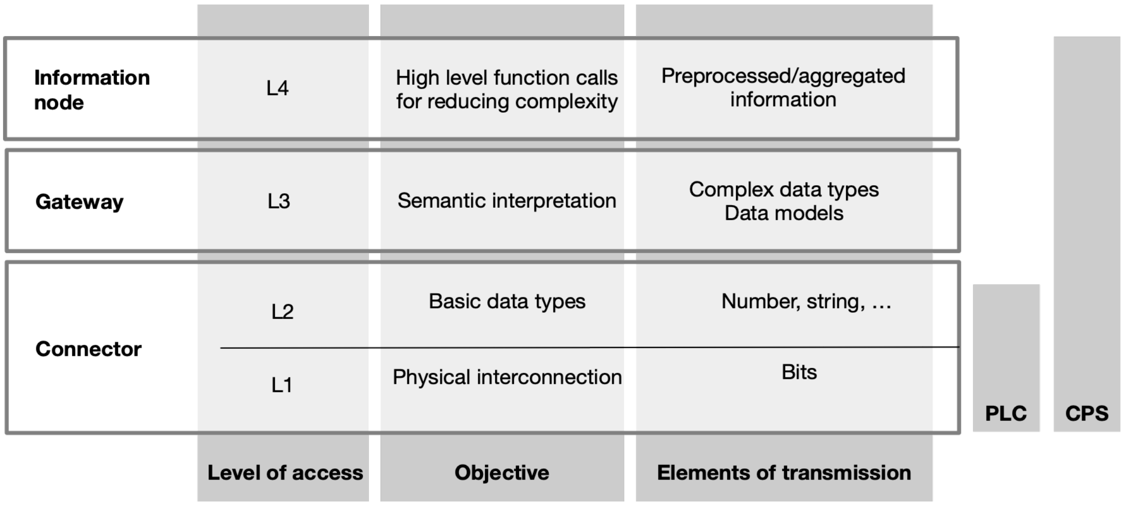


Figure 3: Classification and requirements for an information gateway

Figure 3 shows the requirements for such a component and makes a classification as well as a distinction from typical device classes. An information gateway with firewall function operates at the level of an information node. These capabilities can be realized, for example, as an independent component between production objects (e.g., machine) and platform by an edge controller device (see [20]).

3.2. Reference model

In addition to the hardware and software concepts for an information gateway, further building blocks are required which, as elements of the surrounding overall system, represent the prerequisites for effective deployment. Applying the structure of organization, technology and human, the building blocks shown in Figure 4. As an example, the user access module is detailed here (Figure 5). The data sources (left) and the platform (right) are visible. In between, the information firewall is realized by means of the information gateways and I4.0 box, which regulate the flow of information as a connecting element and in this case provide the necessary functions by means of agents and automate them appropriately.

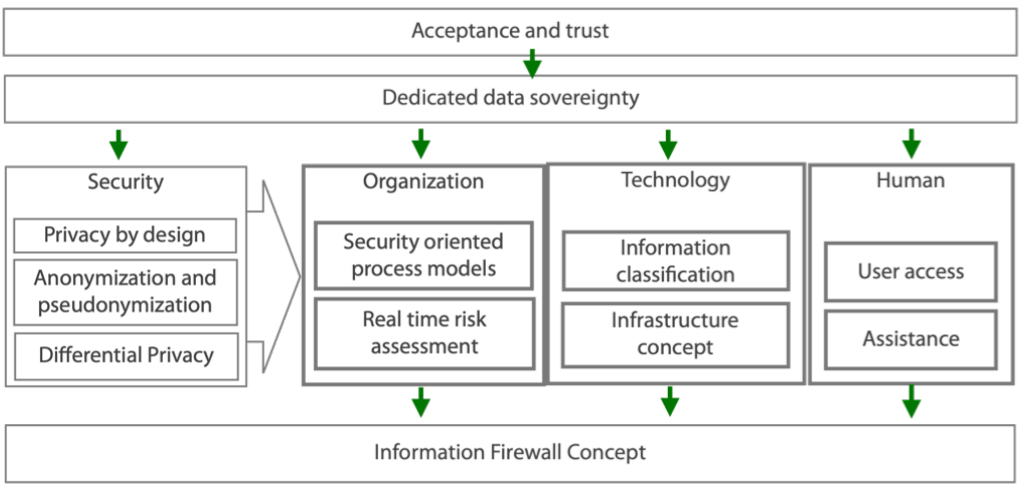


Figure 4: Result for the realization of an information firewall

4. Hardware concept

The deployment of the information firewall concept needs a hardware concept that enables the implementation in existing infrastructures. In reference to the presented building blocks this concept have to fulfill several requirements. Accordingly, potential data sources must be enabled in a proper way to act as CPSs. In particular, from brownfield perspective (see section Influencing factors) the integration of closed legacy systems is a typical use case. The concept envisages a component which enables the required properties to be retrofitted and equips a production object with CPS capabilities. The implementation of the information firewall function follows Industry4.0's ideas of complexity reduction through decentralized elements with capabilities for local information processing. It is therefore obvious to design this component as a typical CPS in accordance with I4.0. Thus, the device acts on the information node level (see section 3.1).

Figure 5: User access with mobile dashboard and I4.0-Box device

The term I4.0 box is used in reference to the Industry 4.0 concept. Also picking up on this aspect, the reference architecture model Industry4.0 (RAMI 4.0) also describes such an approach under the conceptual term management shell [23]. The management shell as an interoperable profile of the technical actor (e.g., the machine) provides information about the data supplied, among other things. Thus, by comparison in combination with the description of the business model, a classification can be made with regard to adaptations of the communicated data. The use of local information processing also contributes to the flat communication hierarchy of IoT. Figure 5 shows a an actual implementation of this hardware concept. CPS enablement, data source with local information processing and communication capability: the device gains access to the production object's installed sensors by connecting them via discrete cabling, using existing fieldbus via the controller/PLC, or alternatively via additional sensors installed at a suitable location. Actuators are connected via fieldbus or via direct cabling to the controller. In analogy to operating systems and their tasks - the abstraction of the operating equipment from the underlying hardware and the management of hardware resources - the software components of the boxes are summarized under the term factory operating system (FaBS). The ConnectionService realizes the communication of the components. Similar to a driver, it abstracts technical details and provides access to FaBS functions. It allows the implementation of a gateway function between the internal communication of the system components and the respective protocol of the external component. This can be done using standards (such as OPC-UA in the case of linking different runtime levels) or using things, specifically defined by the provider.

5. Conclusion

The individual and situational protection needs of data owners are juxtaposed with the need to share information regarding the effectiveness of platform-based business models. By creating dedicated data sovereignty through scalable transparency, this paper addresses this issue and provides an approach to address this problem. It systematically shows the necessary building blocks. The reference model provides a framework for further activities to operationalize the identified building blocks. It offers a basic procedure whose principles serve as a basis for the development of such solutions. An obvious and further approach is to establish the analysis of information flows as an important part of securing the IT infrastructure and to implement the detection of dynamic structures as well as the detection of irregularities and anomalies in suitable systems. To adequately master these tasks, AI-based methods are possible tools for technical operationalization, e.g., to design self-learning (security) systems. The CPS-based hardware concept offers a usable runtime environment in this respect.

Acknowledgements

The IGF project 20664 BG "IoT-Business Model Evolution - Development of a stage-oriented IoT strategy for SMEs in the injection molding industry for the establishment of interoperable platform ecosystems" is funded by the German Federal Ministry for Economic Affairs and Energy via the AiF as part of the program for the promotion of joint industrial research (IGF) on the basis of a resolution of the German parliament.

References

[1] IDG, 2018, Internet of Things 2018. Key findings, IDG Business Media GmbH

[2] Bender, B., Lass, S., Habib, N. et al., 2020, Platform deployment strategies in mechanical and plant engineering: strategies of German companies in the Industry4.0 context. HMD (2020). DOI: 10.1365/s40702-020-00648-1.

[3] Dremel C., Herterich M., 2018, Digital cloud platforms as enablers for the analytical use of operational product data in mechanical and plant engineering. InCloud Computing 2018 (pp. 73-88). Springer Vieweg, Wiesbaden.

[4] Naderer, B., (ed.), 2017, Latest news from kunststoffland NRW e.V. Focus: Opportunities of digitalization (Issue 2), https://www.kunststoffland-nrw.de/download/publikationen/nrw_report_2017-02- screen_5mb.pdf, last checked on 16.04.2021.

[5] Vassiliadis, M. (ed.), 2017, Digitalization and Industry 4.0. Technology alone is not enough. Hannover: Industrygewerkschaft Bergbau Chemie Energie.

[6] Velosa, A., Schulte, W, R., Lheureux, B., 2020, Hype Cycle for the Internet of Things, Gartner.

[7] Gao, L. ,2019, Exploring the Data Processing Practices of Cloud ERP - A Case Study. Journal of Emerging Technologies in Accounting, 17(1):63–70.

[8] VDI 2014, Industry 4.0-On the way to a reference model. VDI Status Report.

[9] Bauernhansl, T., ten Hompel, M.,; Vogel-Heuser, B., 2014, Industry 4.0 in production, automation and logistics: application, technologies, migration: Springer-Verlag Berlin Heidelberg.

[10] Bauer, W., Schlund, S., Marrenbach, D., Ganschar, O., 2014, Industry 4.0 - Economic potential for Germany. Ed. by BITKOM and Frauenhofer IAO.

[11] Vogel-Heuser, B., Bauernhansl, T., ten Hompel, M., 2017, Handbook Industry 4.0 Bd. 4: Springer.

[12] Hasan, H. and Mohammed, B. (2018). Evaluation of MQTT Protocol for IoT Based Industrial Automation. International Journal of Engineering Science, 8:19364–19369.

[13] Theuer, H., 2019, Market overview of IoT gateways. Factory Software, 24(3):58-62, Gito Berlin.

[14] Krause et al, 2017, IT platforms for the internet of things (IoT), IT platforms for the Internet of Things (IoT): Basis of intelligent products and services. Stuttgart: Fraunhofer Verlag.

[15] Andreev, S., Koucheryavy, Y., Baladin, S., (ed.), 2012, 12th International Conference NEW2AN. Springer.

[16] Sethi, P., Sarangi, S. R., 2017, Internet of things: architectures, protocols, and applications. In: Journal of Electrical and Computer Engineering 2017.

[17] Vogt, A., 2017, IoT Vendor Benchmark 2017 Germany: Experton Group AG.

[18] IDG Research Services 2016, IDG Business Media GmbH.

[19] BMWi, 2016, Announcement for the funding initiative Mittelstand 4.0 - further competence centers for innovative solutions for the digitalization and networking of the economy. BAnz AT 26.08.2016: Federal Ministry for Economic Affairs and Energy.

[20] Lass, S., 2017, Benefit validation of cyber-physical systems in complex factory environments - A hybrid simulation concept for Industry 4.0. GITO Berlin.

[21] Strauß et al., 2018, Enabling of predictive maintenance in the brownfield through low-cost sensors, an iiot-architecture and machine learning. In 2018 IEEE International conference on big data (big data), pages 1474–1483. IEEE.

[22] Gronau, N., 2014, The influence of cyber-physical systems on the design of production systems. In Kersten, W., Koller, H., and Lödding, H., editors, Industry 4.0 - How intelligent networking and cognitive systems are changing the way we work, p. 279-295. GITO Berlin.

[23] Adolphs, P., Epple, U., 2015, Status Report Reference Architecture Industry 4.0 (RAMI4. 0).

[24] Lass, S. and Gronau, N., 2020, A factory operating system for extending existing factories to industry 4.0. Computers in Industry, 115:103128.

[25] Eylers, K., 2020, Proposal for the systematic classification of interactions in Industry 4.0 systems. Technical report, German Association for Information Technology, Telecommunications and New Media.

[26] Herzog et al. 2016. Semantic interoperability in IOT-based automation infrastructures: How reference architectures address semantic interoperability. at-Automatisierungstechnik, 64(9):742–749.

[27] Stahl, J.: Accepting Change, Effecting Change-The Role of Employees in Strategic Renewal. In: Krüger, W. (Ed.) ; Bach, N. (Ed.): Excellence in Change. Springer, 2014, pp. 129-161.

[28] Gebauer, L. et al., 2012, Barriers to the use of cloud computing in B2B and the allocation of these to the different trust relationships. In: ConLife 2012 Academic Conference, Cologne, Germany.

[29] Repschlaeger J, Wind S, Zarnekow R, Turowski K., 2012, A reference guide to cloud computing dimensions: infrastructure as a service classification framework. In; 45th Hawaii International Conference on System Sciences 2012 Jan 4 (pp. 2178-2188). IEEE.

[30] Lass, S. and Fuhr, D., 2013, IT-Sicherheit in der Fabrik. Productivity Management, 18(4):29–32.

[31] Wirsam, J., 2008, Know-how as an object of protection in the context of innovation management. In: Spectrum of Production and Innovation Management: Springer, pp. 233- 242.

[32] Bender B, Gronau N., 2017, Coring on Digital Platforms-Fundamentals and Examples from the Mobile Device Sector. InICIS 2017.

[33] Gronau, N., 2014, Adaptability in Production and Logistics. In: Productivity Management 19 (2014), No. 2, pp. 23-26.

[34] Westkämper, E., Zahn, E., 2009, Transformable Production Companies: The Stuttgart enterprise model. Springer, 2009.

[35] Kaufmann, T. and Servatius, H.-G., 2020, The Internet of Things and Artificial Intelligence as Game Changers. Springer. S. 105.

[36] Wadhwa, P., Puri, A., 2016, Internet of Things: Challenges and impact. International Journal of Engineering Research and General Science, 4(3), 781-786. http://pnrsolution.org/Datacenter/Vol4/Issue3/109.pdf:

[37] GAIA-X, 2020. Driver of digital innovation in Europe - Featuring the next generation of data infrastructure Federal Ministry for Economic Affairs and Energy (BMWi) 2020,

[38] Steinbach, M. et al., 2016, Data Protection and Data Analytics: Challenges and Approaches. In: Data Protection and Data Security, Issue 7, 2016; pp. 440 - 445.

[39] Marnau, N., 2016, Anonymization, pseudonymization and transparency for Big Data. In: Data Protection and Data Security, Issue 7, 2016; pp. 428 - 433.

[40] Ulbricht, C., 2015, Anonymization and pseudonymization; encryption. In: Dorschel, Joachim (Ed.), Handbook of Practice Big Data. Springer Fachmedien Wiesbaden, pp. 185 - 189.

[41] Jain, P. et al., 2016, Big data privacy: a technological perspective and review. In: Journal of Big Data, Heft 1, 2016; S. 1 – 25.

Biography

Sander Lass, Dr. (*1976) studied at the Technical University of Berlin and received his doctorate in applied computer science from the University of Potsdam. He conducts post-doctoral research in the field of factory software and is the technical director of the Center Industry 4.0 Potsdam. The focus of his work is the transfer of Industry4.0 building blocks and concepts into practical applications.

Benedict Bender, Dr. (*1989) studied business informatics at the University of Potsdam, the Humboldt University of Berlin as well as the University of St. Gallen. His research interests include Industry 4.0 and aspects of IT security and privacy. Furthermore, he deals with digital platforms and business ecosystems.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Derivation Of Counter-Measures For Industry 4.0 Environments And Cyber-Physical Production Systems Based On Their Cyber-Security Vulnerabilities

Günther Schuh[1], Jacques Engländer*[1], Lars Kaminski[1], Jan Hicking[1], Martin Bülskämper[1]

[1]Institute for Industrial Management, FIR at RWTH Aachen University, Campus-Boulveard 55, Aachen 52074, Germany | *Corresponding author

Abstract

The digital transformation is changing the way companies think and design their manufacturing environment. Both due to the increasing number of connections between IoT-Devices, tooling machines, and production lines and the phenomenon of the convergence of IT and OT, systems are becoming more complex than years ago. Organizational and cultural changes within manufacturing companies strengthen this trend and form Industry 4.0 environments and cyber-physical production systems (CPPS). As these systems do not longer stay alone but are connected to each other and the company's outside, the size of the potential attack surface is increasing as well. Besides that, manufacturing companies, small and medium-sized in particular, are facing complex challenges based on lack of knowledge, budget, and time to understand as well as to interpret their current situation and risk level and therefore to derive necessary counter-measures. Efficient as well as pragmatic tools and methods for these companies do not exist. This paper shows a research approach in which the company-specific set-up of Industry 4.0 environment and CPPS is characterized by its potential vulnerabilities. This enables companies to evaluate their risk potential before setting up this kind of environments and to undJo,erstand the potential consequences more precisely. By doing so, companies can derive and prioritize important counter-measures and so to strengthen their level of cyber-security efficiently. This will decrease the number of cyber-security attacks and increase the company's competitiveness.

Keywords

Cyber-Security; Industry 4.0 Environments; Manufacturing Companies; SMEs; Digital Transformation;

1. Introduction

The digital transformation of manufacturing companies is changing the way on how we think about organizational, technical, cultural, and processual relations. We see shifting paradigms in manufacturing companies through Industry 4.0 [1], which need to be addressed in several ways. These paradigms are required for so-called cyber-physical production systems (CPPS), which are formed by former standalone cyber-physical systems (CPS). Those CPS are physical systems such as machines, tools, or electrical components, enriched with cyber aspects such as software components. Therefore, they are characterized by a deep linking of physical and computational elements. On a higher enterprise level, this convergence of information technology (IT) and operational technology (OT) is called IT/OT-convergence, forming the Industrial Internet of Things (IIoT). It enables companies to now understand patterns, which had been an unknown unknown in the past, by analyzing data from former silos.

The result is, that the security goals of both IT and OT are merging [2]. Due to the fact, that these IIoT-systems are complex by their nature and are becoming even more complex because of new technologies and further developments, manufacturing companies seem to be overburdened by questions such as how to manage and how to secure these systems. As cyber- and information-security have always been an important aspect to take into account, its importance is continuously growing due to the convergence of different domains. It is no longer acceptable just to secure single components, the realization of cyber-security needs to happen on an entire system level [3]. Therefore, the trustworthiness of IIoT-systems is becoming an important success factor for manufacturing companies. While many companies understand the need, they face significant challenges in implementing it. Lots of existing regulations, standards, and frameworks are too extensive to be applicable by SMEs. Furthermore, they are characterized by a lack of brownfield approaches, because the large number is not taking existing frame conditions into account. The fact that SMEs are suffering from limited resources, budget, know-how, and time complicates their situation [4]. Therefore, we focus on the following three aspects in this paper:

- Modeling of Industry 4.0 environments based on existing but adopted frameworks
- Mapping of the models' intersections and most relevant vulnerabilities for SMEs
- Derivation of counter-measures for a cost-benefit-efficient perspective

As these three aspects do also represent a stepwise approach that can be integrated into an overall risk assessment, this paper will help SMEs to identify their largest gaps in cyber-security based on their existing environment. The mapping of vulnerabilities to the Industry 4.0 environment will enable companies to derive them without specific knowledge about cyber-security, but instead about their IIoT-systems. The goal is to show SMEs an efficient and pragmatic way to start first cyber-security initiatives on their own. The scope is not to make norms and standards obsolete, but to lower the entry threshold for SMEs by shifting the focus and starting point to a well-known area, their IIoT-Systems, and Industry 4.0 environment.

2. State of the art

The research in the field of cyber-security has increased over the last years. Especially the combination of cyber-security aspects in manufacturing companies implementing Industry 4.0 use cases has become an important research topic. This section, therefore, presents important norms and standards, IIoT reference architectures as well as approaches for threat modeling and vulnerability identification and shows their gap in comparison to the practical needs of SMEs.

2.1 Norms and standards

Norms and standards related to information security are often designed based on different target groups and fields of topics. They help to increase the information security level as well as to simplify the communication between different institutions about the controls to be applied. [5] As there are different standards across different countries, institutions, and sectors[1], lately, activities tried to consolidate them as well as make them interoperable. Currently, the standard ISO/IEC 62443 is the most actively processed one in the field of Industry 4.0, while it makes use of existing norms and standards [7].

Within the standard ISO/IEC 62443, the International Electrotechnical Commission (IEC) defines a security standard for an industrial automation and control system (IACS). It consists of several parts and subparts. The first part describes foundational terminology and basic knowledge. The second part especially deals with requirements related to an IACS information security management system. Besides that, it consists of methods related to patch management. Security technologies, levels, and requirements in general and related

[1] The European Union Agency For Network And Information Security (ENISA) published a list with around 50 existing information security and privacy standards for SMEs [6].

to specific topics are described in part three. In more detail, part four lastly deals with security requirements for IACS components and general requirements along the product development process of those components. It takes different principles into account, e.g. defense in depth (different security levels arranged from the outside (lowest security level) to the inside (highest security level)) and zones & conduits (system divided into several zones related to specific security levels and the communication in between).

The international and cross-sector standard ISO/IEC 27001 aims to support companies to manage their information security. It specifies requirements for establishing, implementing, maintaining, and continually improving an information security management system (ISMS). Furthermore, the assessment and treatment of risks related to information security is also part of the standard. The way the requirements are set out in ISO/IEC 27001:2013 is generic so that they can be applied to all organizations. [6] The standard is a relevant part of the superordinate row of ISO/IEC 2700X standards, which in total includes over 20 norms.

The Bundesamt für Sicherheit in der Informationstechnik (BSI) in Germany developed and continually improves the IT-Grundschutz. Since 2005 it is divided into the approach for IT-Grundschutz and several IT-Grundschutz catalogs. Besides management aspects, technical, organizational, personal, and infrastructural controls are described. The four norms BSI-Standard 200-1, 200-2, 200-3, and 100-4 are related to requirements to ISMS, the approach to use the standard, risk analysis based on the IT-Grundschutz, and incident management. The BSI-Standard 200-1 is widely comparable with the ISO/IEC 27001 norm and also considers aspects of ISO/IEC 27000 and 27002. Taken every part into account, the standards consist of more than 5000 pages and 1200 controls. [8]

2.2 IIoT Reference Architectures

Reference architectures are important for describing specific systems. Related to IIoT-systems and Industry 4.0 environments, several well-known architectures have been developed over the last decade. Studies with up to 430 investigated papers show that a large number of architectures exist just in the field of Industry 4.0. In the following, we will describe the most important and near-standard reference architectures in more detail.

The Reference Architecture Model Industrie 4.0 (**RAMI 4.0**), formulated in the norm DIN SPEC 9134, represents one of the first reference models related to Industry 4.0. Besides the definition of terms, it describes different assets in Industry 4.0 and their characteristics. In general, RAMI 4.0 differentiates between the physical world (real assets) and the digital world (data and information). Systems are then described in three axes: **life cycle & value stream** (development to maintenance), **hierarchy levels** (product to connected world), and **architectural layers** (asset to business). The structural design of IIoT-systems, assets, and combination of assets are defined with the help of those axes and layers. Each layer is seen alone but interconnected with the layer above and below. [9]

The Industrial Internet Reference Architecture (**IIRA**) is similar to RAMI 4.0 while describing different layers. In contrast to the three-axes architecture of RAMI 4.0, IIRA uses a **viewpoint-**, **concern-** and **stakeholder-approach** to describe IIoT-systems [10]. Especially the viewpoints (business, usage, functional, and implementation) allow the user to extensively describe the system. They provide stakeholders with the necessary detail to describe their concerns.

Within the Norm **DIN EN 62264-1** a process-driven approach for integrating the corporate and control level is focused. It is based on **ISA-95** specifications and takes the **PERA** (Purdue Enterprise Reference Architecture) model as a basis. The norm serves as a relevant standard for defining the automation pyramid. The automation pyramid has been an important model to describe technical systems and their interconnections related to the shop floor and the automation environment. Nevertheless, due to the paradigm shift through Industry 4.0, the pyramid is becoming outdated, even if its generic structure is still applicable to most manufacturing companies.

2.3 Threat Modeling

SMEs facing challenges with their cyber-security often see themselves struggling with deriving the most relevant threats and vulnerabilities concerning their Industry 4.0 environment. Besides existing studies on the most relevant vulnerabilities such as missing employee awareness or missing patch management, external consultancies assessing these environments as well as expensive hardware or software solutions, SMEs seem to stand alone without any guidance on the topic. A possible way for SMEs to derive their vulnerabilities is given by threat modeling. As a process, it represents a special form of model building. By developing a threat model, companies can derive potential threats for the model of the systems concerned [11]. Extensive threat models are often developed and applied as part of system (software) development to strengthen the principle of "security by design". Therefore, the applicability for manufacturing SMEs is limited, even though principles can be valuable guidelines during development [12]. The framework for modeling threats often consists of the following four steps [11]:

1. Modeling of the system to be built, deployed, or changed
2. Finding threats using existing models such as STRIDE
3. Addressing threats based on the model
4. Validating for completeness and effectiveness

It is important to understand that there is not the “one correct” way for modeling threats. The four-step approach just gives the frame conditions which have to be detailed and tailored to the specific situation in which a company is situated. Furthermore, it is important to always focus on the result to be achieved. If the model shows several threats and works to deal with them, it will empower the company. [11] In combination with other tools, threat modeling gains even more power. Existing vulnerability ontologies [13–15], specialized methods for vulnerability identification [16,17], methods like STRIDE [18] or CVE [19], and automated solutions such as AI-based Twitter searches [20] help to increase the benefit of threat modeling.

Therefore, threat models can be used for developing generic frameworks. In this paper, we made use of threat modeling for the development of the framework and approach in chapter 3. STRIDE will serve as the most mature model in this context.

2.4 Research gap

It can be stated that a lot of standards do exist, but even if they are intended to be universally applicable, they are often not applicable to SMEs due to their scope and size [4]. Not least because of this, there are numerous guides for implementing the various standards (see for example [21,22]). SMEs are facing different challenges than bigger companies, such as restricted budget, a lack of knowledge, and missing structures [6]. Same counts for existing architectures. They are mostly not applicable for SMEs to their cyber-security concerns [12]. Especially RAMI 4.0 and IIRA are not useful in a practical way, because of their generic character [23]. Threat modeling and existing threat models give a good starting point for SMEs to start with their cyber-security. But without knowledge within the company and guidance on choosing the right additional tool, it is hard for SMEs to do it with high completeness and effectiveness [24].

3. Framework

In this paper, we propose a framework based on existing standards, reference architectures, and approaches applicable for SMEs. The framework aims to provide SMEs with an instrument to manage their cyber-security related to Industry 4.0 environments on their own and without deep expertise and a high budget. Within the framework, not only technical but also organizational and personal aspects are taken into account to serve as a holistic approach in SMEs. Furthermore, the framework is developed to be applicable for effectiveness and efficiency.

3.1 Structure of the framework

Following KAMAL [25] respectively ZACHMAN [26], a systematic approach to study cyber-security is based on the decomposition of the system to be modeled answering the following questions:

- **What?** – Description of the system and its components relevant to the threat model
- **Why?** – Description of the primary business objective to accomplish a primary goal
- **How?** – Description of the function to achieve the goal
- **Who?** – Description of the human factor in each system
- **When?** – Description of the system's lifecycle
- **Where?** – Description of the components in the context of the physical environment

Those questions can be mapped on a security-related Industry 4.0 reference architecture developed for this paper shown in Figure 1. The '**What?**' represents affected assets through component layers in accordance to e.g. RAMI 4.0. Furthermore, it is extended by the three important aspects Business ('**Why?**'), Function ('**How?**') and Human ('**Who?**'). The '**When?**' represents the security life-cycle, starting with the initiation phase and ending with disposal, to include principles such as security-by-design and phase-specific vulnerabilities. Lastly, the '**Where?**' shows the physical location respectively zones for (sub-)systems as it does the PERA model for example.

The model claims to be comprehensive, layered, and modular at the same time [25]. Comprehensive means that the model offers the opportunity to derive the most relevant threats respectively vulnerabilities, even if it just represents a part of the real world. While looking at SMEs, focusing on the most relevant vulnerabilities instead of filling lots of libraries, helps to stay focused while having a lack of budget. The layered approach helps to reduce complexity within systems, so that companies, especially SMEs, can partially describe their Industry 4.0 environment. Model modularity means that different situations can be described separately from each other, which lowers the entry point for SMEs using the framework. The six questions represent six viewpoints, described below.

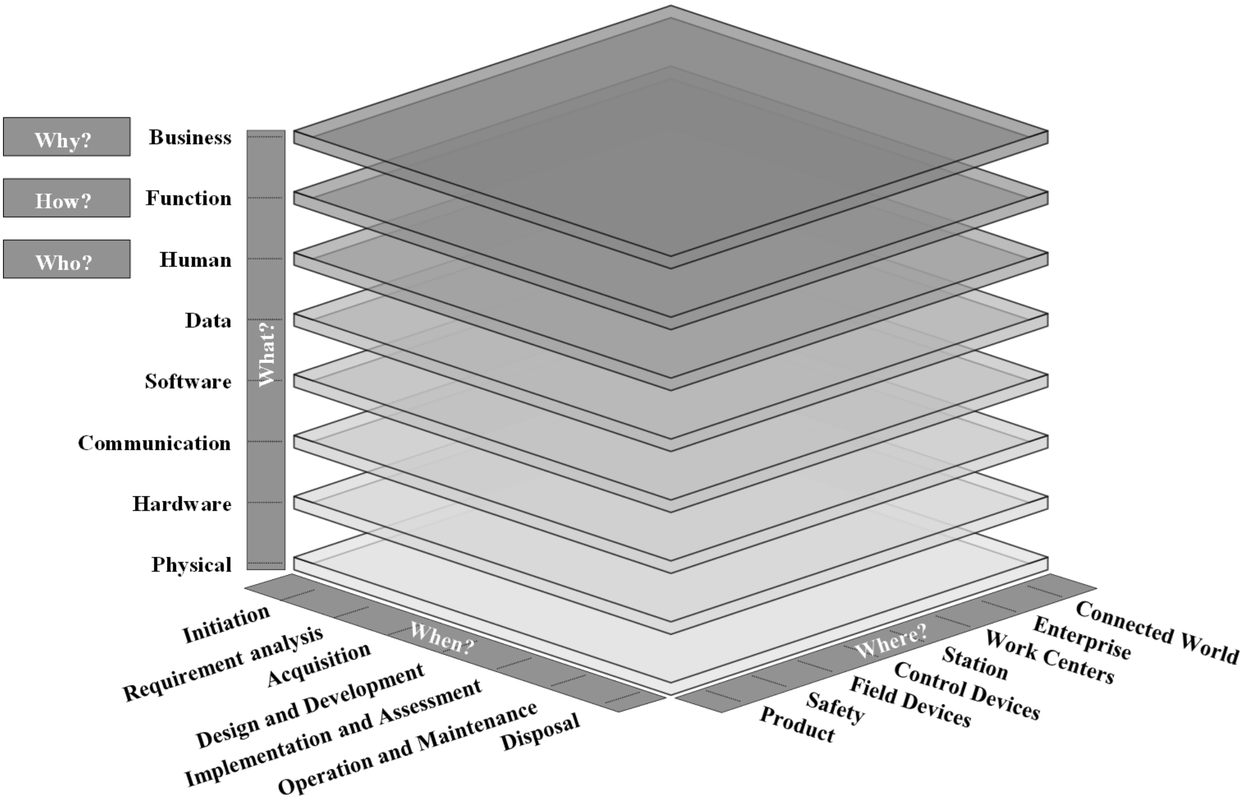


Figure 1: Security-related Industry 4.0 reference architecture

What – Asset-driven: The starting point of modeling the IIoT-system is to make sure to only take those assets into account which are worth protecting the most. These assets can be physical (e.g. machines), software or hardware assets, communication-related assets such as a router, or data in particular.

Why – Business-driven: To be able to make assumptions on the potential risk an IIoT-system is exposed to, the business-driven viewpoint is important to consider business decisions. Whether an IIoT-system has a supportive function or is highly crucial for the stability of the production process implies the needed security level, even if the vulnerabilities may be the same. Furthermore, the viewpoint allows considering stakeholder-specific, supply-chain-related, and tier-oriented aspects.

How – Function-driven: The goal to be achieved by an IIoT-system relies on the system's function. Meant are functional components, their skills as well as their interconnections and interfaces [27]. Concerning cyber-security, a failure of this function would represent an event, which has to be avoided. On the other hand, the system's function should not be influenced negatively by security counter-measures.

Who – Human-driven: To reflect the fact that a large number of attacks on companies are due to employee misconduct [28], the human-driven viewpoint represents the role of the person in front of the system, following the OSI-model and humorously called Layer-8. Besides that, humans are also seen as assets that need to be considered within threat modeling, e.g. concerning environmental threats.

When – Life-cycle-driven: Most of the IIoT-systems do not represent greenfield-like environments. They are embedded in a specific context and situation. Taking these, costs during the design phase and the resulting decrease in business performance into account, the security life-cycle often does not take precedence the running business [29]. Nevertheless, the consideration of each life-cycle phase will help to identify and mitigate vulnerabilities. Each phase is linked to concrete processes and security controls such as the evaluation of tenders within the acquisition phase.

Where – Location-driven: To specify and locate the concrete vulnerability, it seems obvious to model the explicit assets' location. Therefore, the location-driven viewpoint can be seen as an additional viewpoint regarding the assets and their location within the company. For example, IIoT-devices close to the manufacturing process can be mapped on the field device layer, IIoT-devices communicating with SCADA or historian servers on the control layer.

3.2 Framework approach

As the framework itself does not help to identify the most relevant vulnerabilities within IIoT-systems of manufacturing companies, we developed a method in addition to the framework (see Figure 2). It makes use of existing approaches in the field of threat modeling (e.g. [11]) but extends them with a more detailed view on templates and SME-specific challenges. The scalable method works as guidance, mapping the most relevant and universal vulnerabilities to the framework. These vulnerabilities are pre-defined by the author.

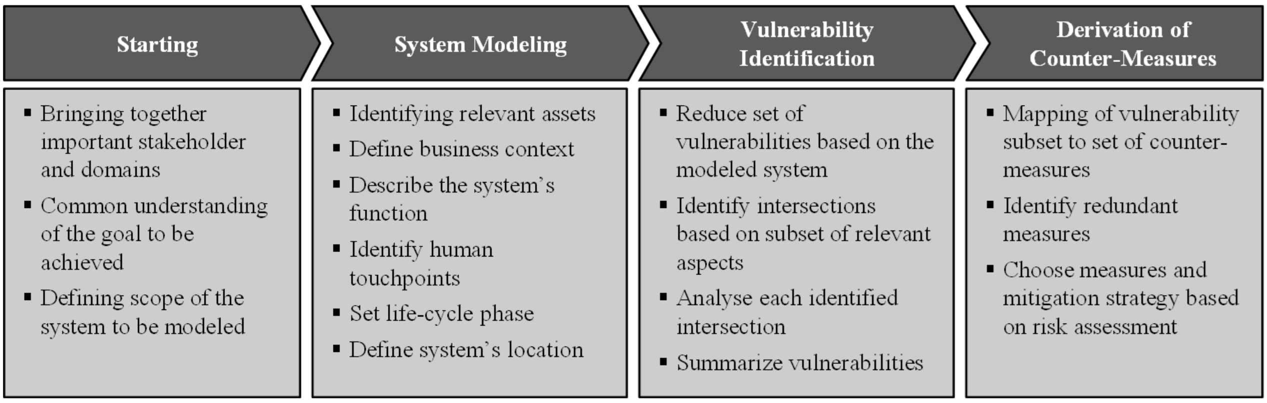


Figure 2: Structure of the approach

First of all, starting a security project needs management attention and important stakeholder from relevant domains at one table. This secures a common understanding of the goal to be achieved by the security project. Additionally, as part of the contextual integration, the common understanding results in condition frames need to be taken into account when defining the scope of the system to be modeled. The distinction between systems and their subsystems is necessary, as this changes elements such as zones and interconnections. Afterward, the before-mentioned questions are based on given templates. These templates are related to existing methods and frameworks, e.g. modeling data flow diagrams or use case diagrams. The application of templates will help SMEs to better understand the connection between their IIoT-system and cyber-security vulnerabilities. In case some assets and their locations, as well as life-cycle phases, are not relevant to the IIoT-system, they are left out. System-related information regarding the business objective will for example decrease the number of vulnerabilities if the IIoT-system does not have interconnections with external entities. After modeling the system, vulnerability identification takes place. The interconnections of each relevant aspect (see Figure 3) will then lead to vulnerabilities linked to those interconnections. Therefore, the result will be a list of relevant vulnerabilities with respect to the modeled systems. Lastly, the derivation of counter-measures is based on a mapping to the identified vulnerabilities. At this point, it is crucial to compare counter-measures among themselves as they may occur more than once. This will lead to the derivation of mitigation strategies which are developed under the premise of cost-benefit-aspects.

4. Use case example

In this section, we use a compromised use case as an example. Figure 3 creates a better understanding of the mentioned intersections. In this case, we look at the intersection of data as an asset, located close to the control device layer within the acquisition phase (**data** × **control device** × **acquisition**). The question to be answered now is which vulnerabilities are linked with this interconnection. In this case, the **STRIDE** framework was used to identify possible vulnerabilities (see Table 1).

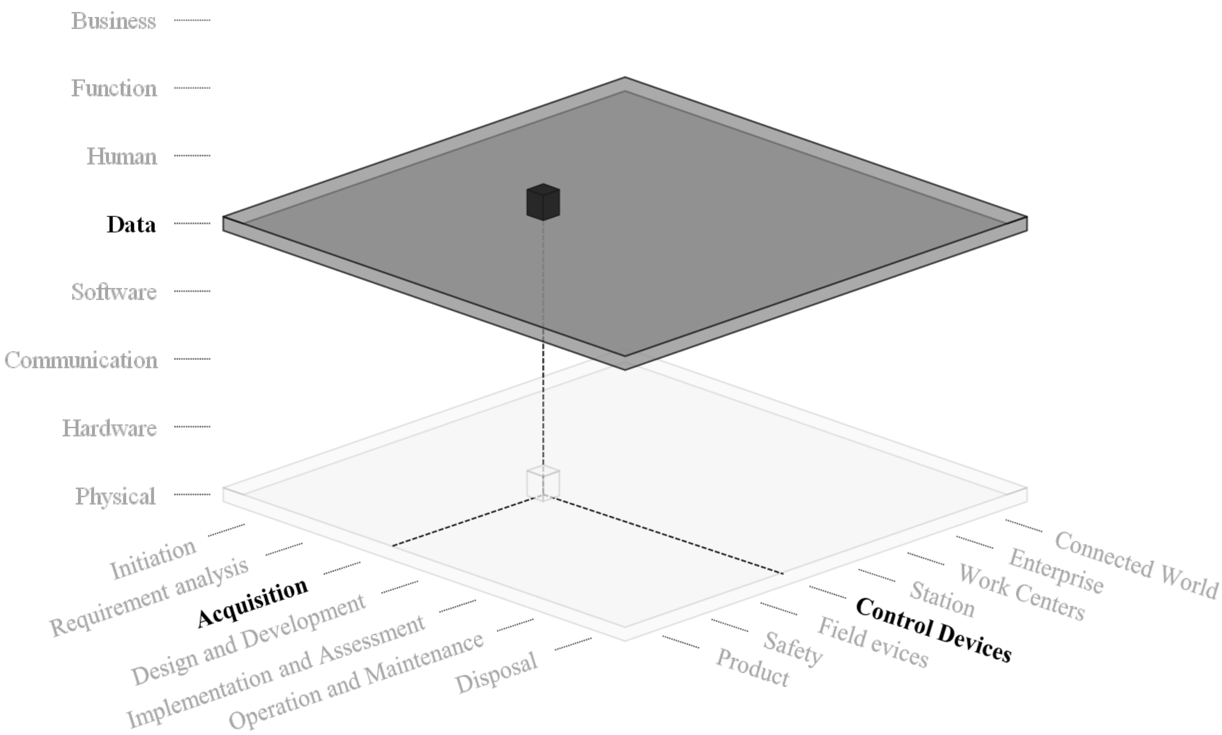


Figure 3: Example of interconnection

In the production environment, there are usually countless control devices that communicate with each other, with field devices, stations, or work centers. It must always be clear who is talking to whom, and that secure authentication is essential. Various authentication methods such as single key, access token, or signatures are possible. Therefore, it is important to think about the appropriate authentication procedure already during the acquisition phase and not to use default keys or methods that are freely available in user instructions or on the internet. This is a widespread attack vector on embedded IoT devices today [30].

Control devices such as PLCs, RTUs, or routers are mainly located in the production environment, where physical access to them cannot be fully controlled. For this reason, it is necessary to consider sufficient physical hardening of the devices already during the acquisition phase to protect their integrity. Physical hardening prevents for example employees or guests from physically manipulating control devices (e.g., drag plugs) which can lead to data loss or an interruption of process control. Furthermore, this can be associated with financial or physical damages, including personal injury.

The purchase of control devices should also be accompanied by the implementation of a suitable event management system. When purchasing control devices, it is therefore important to ensure that they support appropriate functions. This means that access attempts or attacks on individual control devices can be linked to each other (e.g., if IP addresses of accesses are stored). Thereby patterns can be detected, such as attacks on critical production systems. It can also be used to reconstruct and trace attacks.

The acquisition phase deals primarily with the identification of security requirements, the evaluation of proposed security controls, and reviewing and finalizing security design. Therefore, not considering security requirements for control devices such as PLCs represents a vulnerability to the later system, as not assessing the tender and security specification do. Furthermore, this is crucial for future integration activities. [29] For example, the underlying communication of the control devices should be designed with taking future security requirements into account. Otherwise, there is a risk of disclosing confidential data.

A denial-of-service threat affects the availability of devices by using the individual resources of the device. If such attacks are detected at an early stage, appropriate counter-measures can be taken. However, this requires continuous monitoring of the individual data of each device and the search for anomalies. During the acquisition process, care should also be taken to ensure continuous patch management in the future (e.g., that the manufacturer of the control devices provides patches regularly). Among other things, patches are used to eliminate exploits that are used by intruders, for example, to gain additional rights in the system (elevation of privilege). This type of attack is often used, for example, to introduce malware.

Table 1: Exemplary vulnerabilities for data × control device × acquisition

Threat	Property violated	Exemplary vulnerabilities
Spoofing	Authentication	Weak or default authentication methods
Tampering	Integrity	Lack of physical hardening
Repudiation	Non-repudiation	No event management
Information disclosure	Confidentiality	Unencrypted communication
Denial-of-service	Availability	No device monitoring
Elevation of privilege	Authorization	No continuous patch management

5. Discussion and conclusion

The framework and approach presented in this paper show a way to handle cyber-security pragmatically without having profound knowledge or expertise. Therefore, this paper serves as a guide especially for SMEs when facing challenges with cyber-security. Besides numerous other activities, the mentioned approach is an important and necessary step towards SME-specific cyber-security management.

However, as cyber-security is placed in a continuously changing environment, the management of it and the presented method have to be further developed and continuously improved over time. Taking this into account, the method serves as a frame for structuring and handling the topic. Its characteristic of a low threshold, pragmatism, and cost-efficient management address several challenges, but of course does not claim to be complete in a first draft. We consider further activities to be initiated, especially regarding the importance of vulnerabilities concerning the possible risk and its allocation within the model.

References

[1] Schuh, G., 2016. Industrie 4.0 - A Paradigm Shift, in: Haag, C., Niechoj, T. (Eds.), Digital manufacturing. Prospects and challenges, [1. Ed.] ed. Metropolis-Verlag, Marburg, pp. 7–10.

[2] IIC, 2016. Industrial Internet of Things Volume G4: Security Framework. IIC:PUB:G4:V1.0:PB:20160926.

[3] MBMF, 2020. Karliczek: Digitalisierung sicher vorantreiben und technologische Souveränität stärken. Presseportal.de, November 6.

[4] Heidt, M., Gerlach, J.P., Buxmann, P., 2019. Investigating the Security Divide between SME and Large Companies: How SME Characteristics Influence Organizational IT Security Investments. Inf Syst Front 21 (6), 1285–1305.

[5] BSI, 2008. BSI-Standard 100-2: IT-Grundschutz-Vorgehensweise, Bonn.

[6] Manso, C.G., Rekleitis, E., Papazafeiropolous, F., Maritsas, V., 2015. Information security and privacy standards for SMEs: Recommendations.

[7] Ehrlich, M., Wisniewski, L., Trsek, H., Jasperneite, J., 2018. Modelling and automatic mapping of cyber security requirements for industrial applications, in: WFCS 2018. 2018 14th IEEE International Workshop on Factory Communication Systems (WFCS), Imperia. 6/13/2018 - 6/15/2018. IEEE, Piscataway, NJ, pp. 1–9.

[8] Bundesamt für Sicherheit in der Informationstechnik (BSI), 2008. BSI-Standard 100-1: Managementsysteme für Informationssicherheit (ISMS), Bonn.

[9] DIN 91345, 2016. DIN SPEC 91345:2016-04: Referenzarchitekturmodell Industrie 4.0 (RAMI4.0). Beuth Verlag GmbH, Berlin. doi:10.31030/2436156.

[10] Heidrich, M., Leo, J.J., 2016. Industrial Internet of Things: Referenzarchitektur für die Kommunikation. FraunhoferInstiut für Eingebettete Systeme und Kommunikationstechnik ESK.

[11] Shostack, A., 2014. Threat modeling: Designing for security. Wiley, Indianapolis, 590 pp.

[12] Empl, P., Pernul, G., 2021. A flexible Security Analytics Service for the Industrial IoT, in: The Default Password Threat. Virtual, USA. SAT-CPS ´21, ACM, New York, NY, USA. 28. April 2021.

[13] Ahamed, J., Rajan, A.V., 2016. Internet of Things (IoT): Application systems and security vulnerabilities, in: 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA). December 6-8, 2016. IEEE, [Piscataway, NJ], pp. 1–5.

[14] Mozzaquattro, B.A., Agosthino, C., Goncalves, D., et al, 2018. An Ontology-Based Cybersecurity Framework for the Internet of things.

[15] Syed, R., 2020. Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system. Information & Management 57 (6), 1–17.

[16] DeSmit, Z., Elhabashy, A.E., Wells, L.J., Camelio, J.A., 2017. An approach to cyber-physical vulnerability assessment for manufacturing systems. Journal of Manufacturing Systems 43, 339–351.

[17] Hutchins, M.J., Bhinge, R., Micali, M.K., Robinson, S.L., Sutherland, J.W., Dornfeld, D., 2015. Framework for Identifying Cybersecurity Risks in Manufacturing. Procedia Manufacturing 1, 47–63.

[18] Wadhwa, M., 2021. A beginners guide to the STRIDE security threat model. https://www.ockam.io/learn/blog/introduction_to_STRIDE_security_model. Acc. 15 March 2021.

[19] CVE, 2021. Common Vulnerabilites and Exposusres. https://cve.mitre.org/. Accessed 15 March 2021.

[20] Mittal, S., Das, P.K., Mulwad, V., Joshi, A., Finin, T., 2016. CyberTwitter: Using Twitter to generate alerts for cybersecurity threats and vulnerabilities, in: . 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, USA. 18.08.2016 - 21.08.2016. IEEE, pp. 860–867.

[21] ISACA, 2017. Implementation Guideline ISO/IEC 27001:2013: A practical guideline for implementing an ISMS in accordance with the international standard ISO/IEC 27001:2013.

[22] SBS. SME guide for the implementation of ISO/IEC 27001 on information security management. Small Business Standards.

[23] AUTONOMIK Industrie 4.0, 2016. Softwarearchitekturen für Industrie 4.0: RAMI und IIRA aus Sicht der projekte im Technologieprogramm AUTONOMIK für Industrie 4.0. VDI/VDE Institut für Innovation und Technik.

[24] Alshboul, Y., Streff, K., 2015. Analyzing Information Security Model for Small-Medium Sized Businesses, in: 21st Americas Conference on Information Systems (AMCIS 2015). Fajardo, Puerto Rico, 13-15 August 2015. Association for Information Systems, Atlanta, Georgia, pp. 1–9.

[25] Kamal, M., 2019. ICS Layered Threat Modeling. Information Security Reading Room.

[26] Zachman, J.A., 2008. The Concise Definition of The Zachman Framework. zachman.com/16-zachman/the-zachman-framework/35-the-concise-definition. Acc. 17 March 2021.

[27] Fraunhofer ESK, 2016. Industrial Internet of Things: Referenzarchitektur für die Kommunikation.

[28] Zimmermann, S., 2019. Menschliches Fehlverhalten und Sabotage sind die größten Bedrohungen: VDMA Industrial Security. https://industrialsecurity.vdma.org/viewer/-/v2article/render/37164217. Accessed 17 March 2021.

[29] CSA, 2017. Security-by-Design Framework. Version 1.0.

[30] Fraunholz, D., Krohmer, D., Anton, S.D., Dieter Schotten, H. (Eds.), 2017. Investigation of Cyber Crime Conducted by Abusing Weak or Default Passwords with a Medium Interaction Honeypot, 7 pp.

Biography

Prof. Dr.-Ing. Dipl.-Wirt. Ing Günther Schuh is the director of the Institute for Industrial Management (FIR) at RWTH Aachen University. He is also head of the chair of production systems and a member of the directorate of the Laboratory for Machine Tools and Production Engineering (WZL) at RWTH Aachen University as well as if the Fraunhofer Institute for Production Technology IPT in Aachen. His research findings include relevant methods and instruments for complexity management, resource-oriented process cost accounting, and participative change management, as well as the concept of the virtual factory.

Jacques Engländer, M. Sc. RWTH has been working as a project manager at FIR at RWTH Aachen University since 2018. In his current position as part of the Information Management division, he supports companies in the design of IT strategies and information security management concerning organizational, technical, and cultural aspects.

Lars Kaminski, M. Sc. started working as a project manager at FIR at RWTH Aachen University in 2019. In his current position within the technical group IT complexity management, he specialized in IT architecture management and the design of transformation strategies for information security.

Dr. Jan Hicking has been head of the division Information Management at FIR at RWTH Aachen University since 2020. Starting in 2017, he received his Ph.D. in 2020 within the field of intelligent products. As head of the division, he is responsible for multifaceted consulting and research projects.

Martin Bülskämper, B. Sc. RWTH has been a research student assistant at FIR at RWTH Aachen University since 2020. As an automation engineer, he combines technical expertise with management aspects such as the management of information security and thus helps manufacturing companies to successfully manage digital transformation projects.

2nd Conference on Production Systems and Logistics

Federated Machine Learning Architecture for Energy-Efficient Industrial Applications

Can Kaymakci[1,2], Lukas Baur[1], Alexander Sauer[1,2]

[1] *Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Stuttgart, Germany*

[2]*University of Stuttgart, Institute for Energy Efficiency in Production (EEP), Stuttgart, Germany*

Abstract

Due to the rise of new information and communication technologies manufacturing companies have access to huge amounts of power consumption data which are measured by sensors and processed by information systems. One of the most promising applications of extracting value out of the collected data is the detection of anomalies in process data from industrial machines and equipment. Many research and industry use cases apply machine learning (ML) techniques for anomaly detection. These techniques enable manufacturing companies to optimize their manufacturing processes but also to be more energy efficient and therefore have an impact for sustainable manufacturing. Most of the ML applications use central server infrastructures for data collection from different sources to process and analyse it for further usage. Nevertheless, privacy concerns and security risks motivate manufacturers to store the collected sensitive data from the production line locally. Therefore, suppliers of industrial machines (e.g. robots, machine tools) do not have the possibility, to store and analyse the data in the cloud, where data from all the machines of the supplier in different companies could be analysed and used for ML applications. One of the new paradigm shifts in ML is the concept of federated learning (FL) which enables local devices to use ML without sending data to a central server. This paper introduces an architecture for using the concepts of FL in manufacturing processes enabling machine suppliers to use ML for optimizing machine processes in a collaborative manner. Therefore, the more general federated learning concept is extended for industrial machinery and equipment using the industrial communication framework OPC-UA. Our architecture is tested and validated by using an industrial dataset of different compressors' power consumption.

Keywords

Energy Efficiency; Federated Machine Learning; Smart Manufacturing; Anomaly Detection

1. Introduction

Due to the ongoing digitalization of the energy and manufacturing sector, enormous amounts of data and time series from industrial machines and equipment in supply and manufacturing systems are collected each second by sensors and actuators [1]. The possibilities for increasing resource and energy efficiency by using the collected energy consumption data is enormously high. Detecting irregularities or anomalies in the power consumption of industrial machining and equipment with machine learning (ML) models is one of the keys for energy-efficient manufacturing [2]. Therefore, the use of machine learning (ML) approaches enables manufacturing companies to identify unusual or inefficient behaviour in energy consumption and to find the causes of inefficiencies in industrial machining and equipment [3]. This means that anomalies which lead to higher power consumption than necessary, such as those that occur in poorly maintained, outdated or incorrectly controlled systems, can be detected and measures can be taken [4]. Normally, the huge amount of data is stored in a data warehouse or in data lakes and is aggregated and processed for the specific use case or service [5]. Consequently, machine and operating data must be collected decentrally but managed and processed centrally. This requires a high degree of trust, as possible security gaps in a system can lead to information leaks. In the case of industrial data, potential attackers are able to tap into information about manufacturing processes or even problems at individual companies [6]. Highly sensitive process data can lead to significant competitive disadvantages and cause considerable damage [7]. Thus, the tendency of manufacturing companies is to use only local computing capacities for the collection, management and analysis of data [8]. For machine suppliers applying industrial communication standards such as OPC-UA [9] to process data and deploy ML applications, the sensitivity and purely local management of the data can be a serious problem. Data-based services such as anomaly detection by training a ML model with data from different machines are no longer possible when data is managed locally, as local datasets are insufficient for a ML model due to the lack of generalization and the small amounts of local data available. However, the concept of federated learning has been developed that allows models to be trained without sharing data. Therefore, local devices on the field level train and test ML models locally, but exchange the implicit knowledge as model parameters for the purpose of generalization and robustness of ML models for a specific use case [10]. Recent works focus on the overall architecture of a FL system for industrial applications [11]. Nevertheless, a practice-oriented architecture for training and testing ML models for a network of industrial machinery and equipment to detect anomalies in power consumption is not considered. To close this gap the following two research questions are formulated:

- Which components are needed for training and testing ML models in FL systems in industrial applications?
- How can industrial communication standards be used for privacy-preserving anomaly detection in industrial applications?

This article is structured as follows: The second section attempts to define the concepts of anomaly detection and federated learning. Additionally, the potential for manufacturing companies of using the concepts is elaborated by analysing existing literature and comparing state-of-the-art architectures. In section 3 we introduce our architectural approach for anomaly detection with decentralized data as a federated learning system that enables the use of federated learning by integrating machine learning models with the standardized industrial communication standard OPC-UA. Section 4 demonstrates the approach with a use case of anomaly detection in the power consumption of different compressors in a manufacturing company. We discuss and validate the approach with findings from the use case before we conclude in the final section 5.

2. State of the Art

Anomaly detection in energy consumption of industrial applications is mostly used for detecting inefficiencies [2]. The main challenge is detecting increases or decreases in energy consumption not related to the operation mode of industrial applications, such as machines or the infrastructure of a manufacturing system. Connected to fault detection, it is necessary to detect anomalies in energy consumption to take measures like maintenance or turn off systems [12]. To analyze the energy consumption of industrial applications a multi-level monitoring and near real-time data processing is suggested in [13]. Himeur et al. [2] classifies the different applications of anomaly detection related to energy consumption in buildings into five different categories - abnormal behavior, faulty appliances, occupancy detection, non-technical loss detection and at-home elderly monitoring. Except the last category, these applications can also be adapted to manufacturing processes and infrastructure such as building or technical systems. Most of the current work focuses on building new algorithms or ML models for anomaly detection but do not consider the integration into a superordinate FL system.

FL systems are classified into three categories – horizontal, vertical and federated transfer learning [14]. The horizontal FL approach is a system where the same feature space is shared but the sample space is different [10]. The typical example for horizontal FL in the manufacturing industry is a device or machine which collects the same process data (e.g. voltage or current) but is used in different companies or for different applications. The first horizontal FL solution was published by McMahan et al. [10] in which a single user of an Android phone updates model parameters locally. The updated parameters are pushed to the central server where a centralized model is trained with the parameters from several data owners. Vertical FL or feature-based FL is defined as a learning method for distributed data which is different in feature space but can have the same sample space. Considering time-series data in manufacturing and their corresponding feature spaces vertical FL can be used for manufacturing systems that have the same sample space (e.g. timestamp) but different machines collecting different data for detecting anomalies of the whole manufacturing process. One of the most promising but not widely used approaches is the combination of FL and transfer learning which is called Federated Transfer Learning (FTL) [14]. The goal of transfer learning is to transfer the knowledge, which is the model in the case of machine learning, from an existing domain to a new domain. Hence, FTL applications have neither the same sample or feature space nor multiple datasets with heterogeneous configurations. One of the first applications of FTL is FedHealth [15] enabling the recognition of activities measured by wearable devices. In [16] a FTL framework for classifying different sources of electroencephalography (EEG) data from brain-computer interfaces is proposed. FTL could also be used in machine tools which are similar such as milling machines or welding equipment from different manufacturers.

For the effective use of the described FL approach in software applications Lo et al. [17] define architectural patterns which have to be considered when building an industrial FL system. The patterns are classified into four different categories – Client Management, Model Management, Model Training and Model Aggregation. The client management category consists of patterns which define the processes and components for data management, client selection and the clustering of clients by grouping devices into user groups (e.g. using a machine tool in different industries). The model management derives from the model registry in the classic machine learning lifecycle. The main part of the model registry is to store and manage the model parameters and the different versions of the model [18]. Furthermore, a model replacement trigger and a deployment sector is necessary to change models when the performance is degrading. Additionally, the model compressor is introduced as an important step before sending model parameters to the clients. The model training and aggregation category consists of a model trainer component, a registry and different aggregation components such as secure, asynchronous and decentralized patterns.

The concept of building a generic architecture of industrial FL systems is extended by Hiessl et al. [11] introducing “assets” that generate data on the shop floor during operation and FL “cohorts” grouping assets from the same asset type. Therefore, the different asset data characteristics influenced by operating and environmental conditions are evaluated. The different FL cohorts can be compared to manufacturing islands where the random noise in data such as in temperature or humidity is the same and is detectable by the model itself. Training rounds of the FL system is only done in a cohort. For machine suppliers the cohort could be also defined customer-centric where each customer, such as an original equipment manufacturer or specific plant, is a FL cohort.

However, Hiessl et al. [11] want to evaluate the architecture by using open source frameworks such as PySyft or TensorFlow Federated. These frameworks do not consider any industrial communication protocols between the different devices or assets. Open standards such as OPC-UA are not integrated into the frameworks for an easy-to-use FL application in manufacturing industry. Furthermore, the architectures and approaches presented in this chapter do not include the model and data tracking capabilities suggested in

[18]. An easy-to-use architecture for training and testing ML models in industrial FL applications for industrial equipment and machinery is still missing.

3. Concept

To overcome the limitations described above a novel industrial architecture for training and evaluating machine learning models in FL systems is introduced where model training can be accomplished without exchanging data from machines that share the same goal such as anomaly detection. The overall goal of the architecture is to describe and organize a FL system by defining the different components, their modules and relations. It can be considered as a plan for implementing a software system. Technological aspects such as programming languages are abstracted to generate a concept for different implementation strategies. The main goal of the proposed architecture is to enable machine suppliers for communicating trained ML models with the same feature space but machines in different locations for process optimization by detecting anomalies in energy data such as power, voltage or current. Therefore, the federated principles from [11,17] and the industrial communication protocol OPC-UA [9] is used. The automated data pipeline of collecting, pre-processing, training and testing is done by the local clients which are at the field level. The server coordinate the start and end of the data pipeline. Furthermore, the server is responsible for the aggregation of the model parameters that are approximated in each training round. It is important that the architecture only considers the subsection of training and testing ML model as a main part of a federated AI system for detecting anomalies in manufacturing [18].

The architecture consists of three main functional components – machine trainer, global model coordinator and a model database. We define the **machine trainer** as a functional component on a machine tool which is used in manufacturing processes generating data by sensors. The sensor data is ingested into a predefined data pipeline on the machine trainer. The main goal of the machine trainer is to pre-process data, train and validate a model and send the generated implicit knowledge defined in the model parameters back to the **global model coordinator**. Whereas the machine trainer is at the field level communicating directly to the machining tools the global model coordinator aggregates the model parameters generated by more than one machine trainer. The aggregation of the model parameters is done by using aggregation mechanisms such as proposed in [10]. The history and the different changes of model parameters are tracked in the **model database**. The model database is defined as a graph network showing the relations between model changes and the current use of different models on different clients. Hence, the global model coordinator and the machine trainers can be synchronized and the model database can be seen as the source of finding the best parameters for the next training iteration phase. To show the communication between the functional components and the physical assets the **dataflows** between them is presented as well.

a. Machine Trainer

The machine trainer is the functional component which is executed on a client with processing capabilities such as industrial computers or edge devices on machines. One of its three tasks is to communicate with the global model coordinator to receive events and notify about new model updates. Each incoming notification will either be a train or an evaluation request. For flexibility reasons the client automatically downloads all required model specifications and uses the included model compiler to create a model. The required model specification and the weights can be seen as a recipe that is used for building the model on the local machine without training or building it from scratch. After compiling, its weights are overwritten by the downloaded ones, this will uniquely recreate the stored model in the database. In the case of a train request, this model is used as a pre-trained model and further optimized. For the purpose of evaluation, clients private test data is used to calculate metric scores such as mean absolute error (MAE) for regression tasks or the accuracy for classification tasks. After the evaluation procedure, the optimized model weights and the metric scores are transferred to the model database.

b. Global Model Coordinator

The global model coordinator plays the major role in the communication between the machine trainers and the model database. The global model coordinator hosts an OPC-UA server and implements an interactive terminal for human-interaction with the connected devices and the network. All requests and events are send or received from or to this functional component. It is responsible for creating a base model which serves as the starting model for the first training round. A training round initiated by the global model coordinator is defined as the sum of one training and communication process on each active machine trainer. From the mentioned base model all other models are derived during the iterative training rounds. Also, the coordinator instructs all clients to train on a certain model by updating the respective model OPC-UA node and it collects all client responses to keep track on who is working on which model. Moreover, it receives the model training termination messages from each client after each training round. As soon as all termination notifications for a model are received, the merging phase is started: All respective new model weights are downloaded from the model database and a combined model is calculated by using FederatedAveraging [10]. Finally, the coordinator forwards the evaluation request from a user by triggering an evaluating event. This event starts the evaluation procedure on available machine trainers that test the local data on the merged model. Especially in the case of manufacturing systems the used machines depend on the process. Therefore, the evaluation on different clients is necessary to overcome problems of overfitting and reduce generalization error.

c. Model Database

The database manages four types of components which can also be seen as tables with references to each other: models, structures, model evaluation results and learning graph entities. The former two encode a model uniquely by splitting it into its model structure and model weights. Storing a precompiled model in the database might lead to problems using different client devices and operating systems installed. Therefore only the relevant information such as weight and structure is stored to compile a model from scratch. Whenever a model is needed from the model database the serialized plain description of the model is downloaded and compiled by the machine trainer as a compatible instance. The structure of the model is saved as a reference to look up its structure in the structure table. The model evaluation metrics are uploaded by the clients directly and stored for each model-client pair, since a model performs differently on different local client data. The centralized storage allows the evaluation process to work fully asynchronously. Since in each federated learning training round a model for each client and a combined model resulting from the merge process is created, the number of models grow fast. To keep track of the model evolution, a directed acyclic graph inspired by [19] for model tracking is built up and updated whenever a new model is created.

d. Dataflow

In this section the dataflow (Figure 1) is conceptually explained by drafting a typical learning iteration in more detail. The OPC-UA server on the model coordinator side is stated and at least one client is online. The clients register on the server's model node in their booting code. An initial model separated into its model structure, its initial model weights and its parameters of the optimizer is uniquely defined in the database by storing a foreign structure ID a foreign optimizer-entry ID and its weights. The server starts an iteration round by overwriting the model node by a database model ID. The global model coordinator automatically triggers an event to all subscripted clients. Each available client will immediately respond to the server to await for signals to train on the new global model in this round.

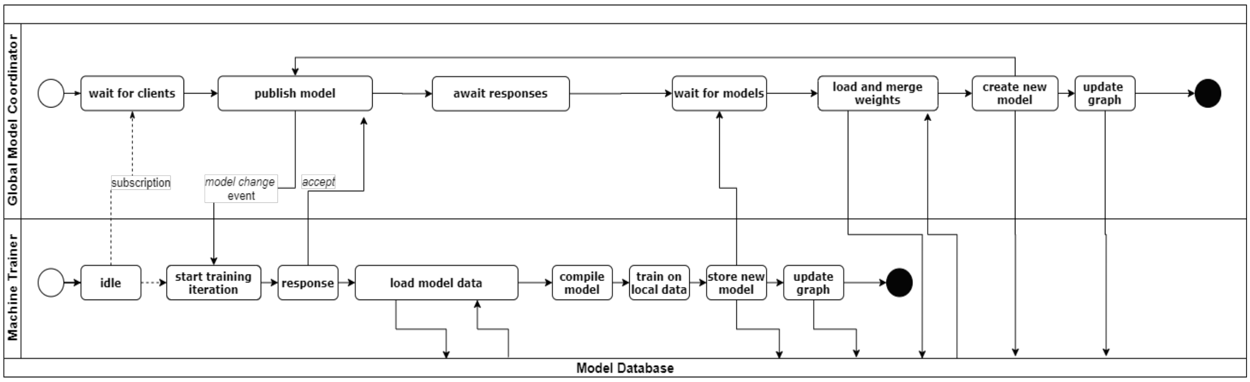


Figure 1: Typical sequence of model training and the interaction between the components

After the registration thread has finished his server response, a model loading thread is started. Given the server's global model ID, the model structure, the model optimizer parameters and the weights are loaded. A parser will compile the model based on the low-level representation. After being successfully instantiated the model, is passed on to a train worker thread performing the actual training on the local data. Once the training has finished, a new model is uploaded to the database uniquely defined by the old global model structure ID, the old model optimizer parameter ID and the new weight vectors. As a final step, the model clients update the training graph database accordingly. It completes its work by sending a response containing the new model ID to the server.

Once all active clients returned their updated model, the server will download the respective weights implicitly linked by the model IDs. A new weighting is calculated by the server using a standard weight averaging approach, similar to [5]. The new global model is stored to the database by uploading the new merged weights and referencing the structure and optimizer. Lastly, the server updates the graph database to keep track of the merging phase. Now the process will start all over again.

OPC-UA intuitively motivates the use of asynchronous working threads. We therefore decoupled the evaluation from the learning phase. We are interested in how the evaluation error evolves over time i.e., how the global model performed on unseen data on various devices and how the learning weights have been distributed during the iterations. Starting from the initial model ID, we scan the graph network for the resulting global models. Having collected all IDs, we trigger the clients asynchronously to start downloading the respective models and evaluate them. After finishing the evaluation, the clients upload their scores to the evaluation database.

4. Use Case

For the purpose of validating the functionality and the convergence of the training phase from machine learning models in a distributed way, a use case from a manufacturing company for anomaly detection is considered. In a manufacturing system different oil-injected screw compressors for further applications such as injection moulding machines are used. The compressors are procured from the same supplier. Furthermore, the compressor types are different, which can be seen in the product data sheets and the overall performance (e.g. working pressure, rotational shaft speed or nominal shaft power). The compressors are installed in a manufacturing plant and integrated into the infrastructure for other machines and systems. For simulating and validating our approach historic data of the power consumption of five compressors is divided into five separated datasets that are transferred on five different simulated edge devices which are considered as virtual compressors. In our experiments, the virtual compressors are simulated on laptop computers with different operating systems and hardware specifications but also single-board computers such as a RaspberryPi. The data is sampled in 15 minutes intervals and pre-processed locally on the edge device. Each client splits its data into 80 percent training and 20 percent testing data und scales the data in a pre-processing step.

By integrating the developed machine trainer into the virtual compressors it is possible to collect and pre-process data but also train and test models without sending the power consumption data. Furthermore, the communication between the virtual clients and the global model coordinator which is deployed on a personal computer is tested and validated. The machine trainer is implemented on all devices by using Python 3.8 and the Python package “freeopcua” to integrate OPC-UA into the environment. The global model coordinator was hosted on one of the virtual compressors. In principal, the model database and the server could have been set up on any other virtual computer accessible in the network. By using different hardware, data of power consumption from oil-injected screw compressors, a closed machine environment and the communication between different compressors for detecting anomalies in their power consumption can be simulated.

The model for detecting anomalies is a LSTM-CNN-based forecast model which uses the predicted power consumption for detecting anomalies. Therefore, the difference between the predicted and the actual consumption is calculated [20]. If it reaches a defined threshold, the power consumption at a given timestamp can be classified as an anomaly. The training and optimization of the Huber loss function is done by the Adam optimizer. For implementing our model we used the Python packages “Keras” and “TensorFlow”. Our case study is divided into two different scenarios. Training a ML model on data from one virtual client is considered as the baseline. For each federated learning experiment, we always performed five *training iterations*. A training iteration is defined as single training on local data of the virtual compressor with a batch of size 32 and 20 learning epochs.

To compare the performances of the federated learning data privacy-preserving approach, we set up two main experiments seen in Figure 2 and 3: In the *reference experiment* we trained a ML model based on the data of a single virtual compressor which can be seen as baseline.

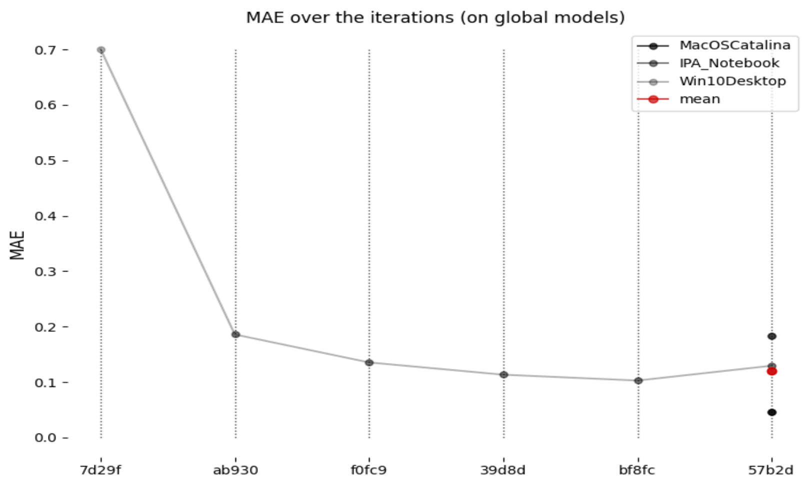


Figure 2: Progress of training in the reference experiment

This is related to a model which is only accessible within a compressor with computing capacity and does not perform any parameter exchange. To argue about its general performance, we evaluated it on both its own test data and the test data of other clients by calculating the mean absolute error (MAE). Consequently, the trained ML model is evaluated by using not only data from the same virtual compressor that the model did not see in training, but also data from different compressor types. Secondly, we performed the *federated learning experiment* using multiple clients to combine their weights after each iteration resulting in a new global model which is used as a starting model for each client in the next iteration. In Figure 2 the decrease of the MAE of the reference experiment can be seen. The x-axis shows the ID of the model which is used for training and testing. The ID of the virtual compressors seen in Figure 2 are the names of the computing devices. fter the complete training phase, the last model is evaluated on two other virtual compressors.

In Figure 3 the federated learning experiment is shown by plotting the MAE after every training iteration and model update. For a better visualization the MAE on the y-axis is plotted in a logarithmic scale. The red line shows the average of the MAE validated on each test data and iteration. The comparison of the two resulting experiment plots show two key insights. Obviously, both experiments converge to some final value, but the convergence speed differs: The evaluation error on the models after the first iteration performs on average better that the reference training. This effect can be measured even better when the number of clients is increased further. After the first two model exchange iterations, the training error remains stable, convergence is approached. Secondly, the weight-averaging pays off in terms of variance reduction: While the MAE reduced in both experiments, the federated learning experiment reduced the overall evaluation error significantly better than the local training (final average MAE of 0.12 for reference, 0.06 for FL experiment). This is because more heterogenic power consumption data from different compressors have been implicitly taken into account, due to the model exchange. Assuming equally powerful computing capacity, a negligible small time overhead for communication and small merging times, due to the parallel training the FL approach results a variance reducing global model in a similar training time. Hence, more power consumption data can be processed and integrated into a combined model for detecting anomalies in power consumption of each compressor. Furthermore, the experiment shows that different machine types of industrial oil-injected screw compressors with the same feature space and target values (detecting anomalies in power consumption) can be integrated into the proposed architecture. The effort for machine suppliers and manufacturers to train and test ML models on different machines decreases significantly.

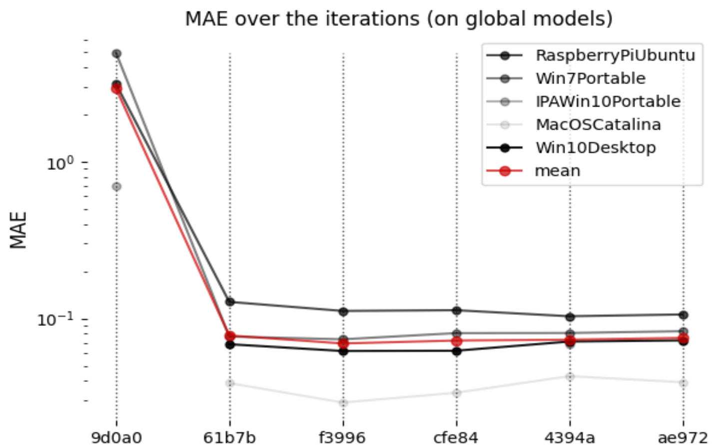


Figure 3: Progress of training in the FL experiment

5. Conclusion

Our presented use case allows the simulation and validation of the training and testing phase of the architecture enabling machine learning for industrial machinery and equipment. There is no need to collect and aggregate sensitive manufacturing data such as power consumption allowing potential attackers to gain information about production time or the usage of machine tools. In our study, we tackled the research questions how to combine the research fields of FL and use cases in manufacturing for anomaly detection into an industrial federated system. Therefore, we developed a generic architecture of communicating model parameters of a machine learning model during the training and testing phase. With our architecture we depict the central components, structure them, and reveal their relationships. Based on the architecture the three important components for training and evaluating ML models in FL for machines used in manufacturing are the machine trainer, the global model coordinator and the model database. Consequently, the architecture and the functional components as research artifacts were prototypically implemented and validated.

Our results have several technical and organizational implications. First, we showed the possibility to combine the actual trends in industrial communication standards such as OPC-UA and privacy-sensitive machine learning environments in Python. Typical libraries (e.g. Keras and TensorFlow) can be used for training local data and aggregating it in a global model. Second, by applying FL manufacturing companies can use machine learning for anomaly detection without losing data ownership and let service providers get access to their data. Last, machine suppliers (e.g. compressor manufacturer) are able to sell machine learning services with guaranteed data ownership to their customers. In addition, our architecture is independent of machine type or operating system and acts as a system for training a model in which different machines participate. This is especially applicable for industrial machinery and equipment with computing capacity (e.g. edge devices) where machine learning can enable optimized processes. To extend the work in this paper the whole machine learning lifecycle and its AI system from data collecting to model tracking as mentioned in [18] will be developed. Therefore, the components and the information flows will be further validated and analyzed to guarantee a continuous improvement process.

References

[1] Chasin, F., Paukstadt, U., Ullmeyer, P., Becker, J., 2020. Creating Value From Energy Data: A Practitioner's Perspective on Data-Driven Smart Energy Business Models. Schmalenbach Bus Rev.

[2] Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A., 2020. Anomaly detection of energy consumption in buildings: A review, current trends and new perspectives, 35 pp. http://arxiv.org/pdf/2010.04560v1.

[3] Chen, H., Fei, X., Wang, S., Lu, X., Jin, G., Li, W., Wu, X., 2014 - 2014. Energy Consumption Data Based Machine Anomaly Detection, in: 2014 Second International Conference on Advanced Cloud and Big Data. 2014 Second International Conference on Advanced Cloud and Big Data (CBD), Huangshan, China. 20.11.2014 - 22.11.2014. IEEE, pp. 136–142.

[4] M. De Nadai, M. van Someren, 2015. Short-term anomaly detection in gas consumption through ARIMA and Artificial Neural Network forecast, in: 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings. 2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings, pp. 250–255.

[5] Patel, P., Ali, M.I., Sheth, A., 2018. From Raw Data to Smart Manufacturing: AI and Semantic Web of Things for Industry 4.0. IEEE Intell. Syst. 33 (4), 79–86.

[6] Mohammed M. Alani, Mohamed Alloghani, 2019. Security Challenges in the Industry 4.0 Era, in: , Industry 4.0 and Engineering for a Sustainable Future. Springer, Cham, pp. 117–136.

[7] Maschler, B., Jazdi, N., Weyrich, M., 2019. Maschinelles Lernen für intelligente Automatisierungssysteme mit dezentraler Datenhaltung am Anwendungsfall Predictive Maintenance. 0083-5560.

[8] Chen, B., Wan, J., Celesti, A., Di Li, Abbas, H., Zhang, Q., 2018. Edge Computing in IoT-Based Manufacturing. IEEE Commun. Mag. 56 (9), 103–109.

[9] Leitner, S.-H., Mahnke, W., 2006. OPC UA - Service-Oriented Architecture for Industrial Applications. ABB Corporate Research Center. https://pi.informatik.uni-siegen.de/gi/stt/26_4/01_Fachgruppenberichte/ORA2006/07_leitner-final.pdf.

[10] McMahan, H.B., Konečný, J., Yu, F.X., Richtárik, P., Suresh, A.T., Bacon, D., 2016. Federated Learning: Strategies for Improving Communication Efficiency. http://arxiv.org/pdf/1610.05492v2.

[11] Hiessl, T., Schall, D., Kemnitz, J., Schulte, S., 2020. Industrial Federated Learning – Requirements and System Design, in: de La Prieta (Ed.), Highlights in Practical Applications of Agents, Multi-Agent Systems, and Trust-worthiness. The PAAMS Collection, vol. 1233. Springer International Publishing, Cham, pp. 42–53.

[12] Henning, S., Hasselbring, W., Burmester, H., Möbius, A., Wojcieszak, M., 2020. Goals and Measures for Analyzing Power Consumption Data in Manufacturing Enterprises, 24 pp. http://arxiv.org/pdf/2009.10369v1.

[13] Henning, S., Hasselbring, W., 2020. Scalable and Reliable Multi-dimensional Sensor Data Aggregation in Data Streaming Architectures. Data-Enabled Discov. Appl. 4 (1).

[14] Aledhari, M., Razzak, R., Parizi, R.M., Saeed, F., 2020. Federated Learning: A Survey on Enabling Technologies, Protocols, and Applications. IEEE Access 8, 140699–140725.

[15] Chen, Y., Wang, J., Yu, C., Gao, W., Qin, X., 2019. FedHealth: A Federated Transfer Learning Framework for Wearable Healthcare. https://arxiv.org/pdf/1907.09173.

[16] Ju, C., Gao, D., Mane, R., Tan, B., Liu, Y., Guan, C., 2020. Federated Transfer Learning for EEG Signal Classification. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference 2020, 3040–3045.

[17] Lo Kit, S., Lu, Q., Wang, C., Paik, H.-Y., Zhu, L., 2020. A Systematic Literature Review on Federated Machine Learning: From A Software Engineering Perspective. http://arxiv.org/pdf/2007.11354v7.

[18] Kaymakci, C., Wenninger, S., Sauer, A., 2021. A Holistic Framework for AI Systems in Industrial Applications, in: Ahlemann, F., Schütte, R., Stieglitz Stefan (Eds.), Wirtschaftsinformatik 2021 Proceedings.

[19] Thulasiraman, K., Swamy, M.N.S., 1992. Graphs: Theory and algorithms. http://onlinelibrary.wiley.com/book/10.1002/9781118033104.

[20] Ullah, W., Ullah, A., Haq, I.U., Muhammad, K., Sajjad, M., Baik, S.W., 2020. CNN features with bi-directional LSTM for real-time anomaly detection in surveillance networks. Multimed Tools Appl, 1–17.

Biography

Can Kaymakci has been working as a research associate in the Industrial Energy Systems Department at Fraunhofer IPA. His current research topic is the concept and development of distributed machine learning applications for manufacturing with federated learning.

Lukas Baur joined the Industrial Energy Systems Department at Fraunhofer IPA as a research assistant in August 2020. Since then, he has been supporting the development of internal machine learning framework components.

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

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Quality Monitoring Of Coupled Digital Twins For Multistep Process Chains Using Bayesian Networks

Maximilian Selch[1], Albrecht Hänel[1], Uwe Frieß[1], Steffen Ihlenfeldt[1,2]

[1] *Fraunhofer Institute for Machine Tools and Forming Technology IWU, Dresden, Germany*
[2] *Technische Universität Dresden, Institute for Machine Tools and Control Engineering, Dresden, Germany*

Abstract

The digital representation of physical assets and process steps by digital twins is key to address pressuring challenges like adaptive manufacturing or customised production. Recent breakthroughs in the field of digital twins and Edge-based AI already enable digital optimization of individual process steps. However, high-value goods typically include multiple step process chains including a broad range from generative and additive processes over several steps of material removal up to assembly. Therefore, a digital twin over the holistic process chain is necessary. While even the set-up of representative twins for a single step is already challenging, a concept for monitoring of the interaction and overall quality control of holistic process chains does not exist yet. The paper introduces a machine-learning method based on probabilistic Bayesian networks to monitor the »digital twin quality« of coupled digital twins which includes several sub-instances of digital twins. The approach identifies the contribution of each instance to the overall prediction quality. Furthermore, it is possible to give a range-estimation for the prediction accuracy of the individual sub-instances. It is therefore possible to identify the most influential sub-instances of digital twins as well as their individual prediction quality. With the help of this information, the quality of the digital twin can be improved by considering individual sub-instances in a targeted manner. Finally, a preview emphasises the potential benefits of the quantum computing technology when dealing with parallel computation of large-scale inference models.

Keywords

Digital twin; Process chain model; Digitalisation; Machine Learning; Bayesian Networks; Quantum Computing

1. Introduction

In manufacturing, production processes generally consist of various production steps. The aim of each single step is to perform an ideal processing after specification of a certain target geometry. Naturally, the practical implementation of a step does not conform the theoretical ideal and shows deviations from the target shape. An important issue is how these deviations of a specific step transmit and affect the quality of succeeding dependent steps within the production chain, including interdependencies between multiple steps [1]. The same applies for digital twins, digital representatives providing all relevant information via a uniform interface, of single production steps: A model of such a specific subsystem is subject to individual uncertainties. Furthermore, the digital coupling of sub-models causes additional uncertainties, which are often non-transparent. To enable a qualitative digital twin over the holistic process chain, it is therefore of great importance to identify the quality of each sub-instance of the holistic digital twin as well as the degree of dependencies between coupled sub-instances. With the help of this information, the strength of the impact that single sub-models have on the holistic digital twin's quality can be estimated. For a probabilistic modelling of the quality of single determining factors and their dependencies on each other, Bayesian networks have proved as an appropriate methodology to reproduce such characteristics of a system [2]. Individual factors can be evaluated by the probability of positive outcomes, e.g. holding a tolerance value as to the target geometry, whereas the impacts on dependent factors are represented by conditional probabilities [3].

After providing an overview of various applications, we classify Bayesian network into the context of manufacturing and refer to some related approaches with Bayesian network in this field of application. A technical chapter then introduces the concept of Bayesian networks in more detail and shows how they can identify and estimate the degree of dependencies. Afterwards, we apply this methodology to a real-world manufacturing setup. We use a single digital twin of a milling process in order to access available real machining data for demonstration purposes, which is directly transmissible to the aspired field of multi-step process interdependencies. Finally, we briefly introduce the concept of quantum computing and depict the potential benefits of this computing technology when working with large-scale Bayesian networks.

2. Related work

Bayesian networks, initially introduced by [3] in 1988, have a broad range of applications in various fields of research as well as industrial settings in the context of modelling uncertainty, risk analysis, decision-making processes and error detection [1,4] covering amongst others the application fields of medicine and healthcare [4] as well as finance [5], supply chain management [6] and predictive maintenance [7]. Furthermore, the usage of Bayesian networks has a significant benefit over other Machine learning methods in relation to the required database since this methodology does not explicitly require massively large-scale datasets and can extract accurate and meaningful statements based on limited data [4]. In manufacturing, it is highly challenging to detect the causes of deviations due to heterogeneous data structures and impacts. Particularly this problem is well resolvable using Bayesian networks.

In this context, Bayesian networks are used as a data analytics tool to reduce system faults as they proofed to predict which components of anonymised measurements in manufacturing datasets will fail at the final stage of the production process [8]. Another use case of this methodology is the improvement of quality consistency in assembly processes where the key is to construct a raw network structure that shows the causal relationships and influence factors on quality consistency, exemplary demonstrated for a diesel engine production line [9]. Aside, [10] considers a general quantification of uncertainty in manufacturing processes focussing on the caused energy consumption as a key factor of sustainability performance. In [11] the technique of inverse inference in manufacturing process chains is demonstrated using a Bayesian network to adjust process configurations in order to achieve certain desired properties, also known as inverse analysis.

In recent times, methods of Bayesian optimization are used in manufacturing background to regulate black box models and to proceed hyperparameter optimization [12,13]. Here we strongly want to distinguish the approach of using Bayesian networks from these considerations.

3. Methodology of Bayesian networks for quality monitoring of digital twins

A Bayesian network is a directed acyclic graph whose nodes represent the influencing factors as random variables and whose edges model the dependencies between the factors, directing from the decisive to the dependent factor. The independent variables of the system hold a probability distribution each which indicate the likelihood of different possible status. The dependent nodes are equipped with conditional probability tables in case of using discrete variables. When dealing with continuous variables, the nodes are mostly assumed as normal random variables. To model connections between multiple nodes, the parameters of the Gaussian distribution representing a child node are dependent on the value of its parents [3]. Since our considerations focus on the strength of influence estimation of certain factors based on different pre-defined status, we motivate the usage of discrete nodes. Of course, if we would focus on representing the very physical process of a specific application with a Bayesian network, the required discretisation of factors having a continuous domain causes approximation errors. Here we want to clarify that this specific task differs from the paper's intention of depicting a quality monitoring framework to handle multiple related digital twins, which may represent different physical processes.

In the application of coupled digital twins, Figure 1 gives an example with various sub-instances of the holistic twin showing probabilities of holding a tolerance value, which influences the quality of the combined modelling. For simplification, the variables are assumed to be binary with the probability of $\boldsymbol{P}(+)$ to fulfil a given tolerance criteria for the modelling quality and a probability of $\boldsymbol{P}(-)$ to violate this criteria. As an example, a deviation from a target position with respect to a certain axis can be considered.

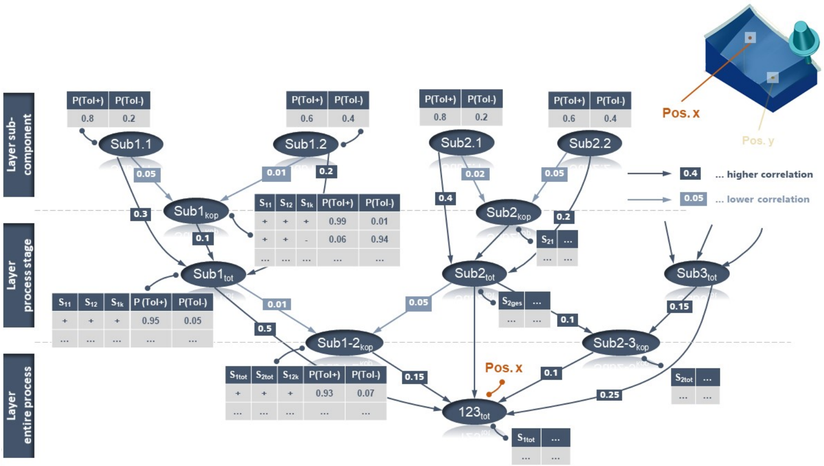


Figure 1: Exemplary Bayesian network for coupled digital twins considering multiple process layers

3.1 Why using Bayesian networks

For the tasks of influence evaluation, deviation analysis and quality improvement of specific machining processes various statistical and ML-based approaches have been proposed. These methods mostly focus on regression and classification models [14], fault tree analysis [15] as well as design of experiment techniques [16]. Often, an advanced adjustment to the considered physical process is required to set up empirical formulas. Bayesian networks allow to make statements on dependencies, strengths of influences and quality without the need of such investigations and enable a comparison between multiple manufacturing processes, which possibly are subject to different physical processes. For such purposes, Bayesian networks especially can cope with highly transient as well as non-linear processes and offer to

- Lean on a limited basis of information,
- Work with uncertain or incomplete data,
- Evaluate the inherent structure of influences,
- Rate the certainty.

Bayesian networks can be build up in three different manners. One way is to use expert knowledge only to construct the network structure, i.e. the dependencies between factors, and define its parameters, namely the (conditional) probabilities of each node. Another approach is to deploy pure datasets without any additional information and learn the network structure as well as the parameters by applying estimation procedures. In between these two ways there is also a hybrid approach using both data and knowledge, which will arise as the appropriate strategy for this paper's application scenario presented in section 4.

3.2 Structure and parameter learning

If the dependencies of the process properties are not specified by domain knowledge, the network structure is learnable either via constraint- or score-based methods. The first set of algorithms use conditional independence test to find the dependencies whereas the score-based approaches produce a series of candidate networks and rank them with correspondent scores [3,17]. Besides, there are hybrid methods like the widely used hill climbing algorithms.

Since the structural dependencies are already defined in the application setup of coupled digital twins (from sub-instances of the holistic twin to coupled models), the key focus is on how to evaluate the degree of these dependencies and to estimate the overall prediction quality. Formally speaking, a Bayesian network with nodes $\boldsymbol{\mathcal{X}} = \{\boldsymbol{X_1}, \ldots, \boldsymbol{X_m}\}$ reflects a unique joint probability distribution $\boldsymbol{\mathcal{P}(\mathcal{X})}$ given by the product of all conditional probability tables

$$\mathcal{P}(\mathcal{X}) = \prod_{i=1}^{m} \mathcal{P}\big(X_i \big| \mathrm{Pa}(X_i)\big) \tag{1}$$

where $\mathrm{Pa}(\boldsymbol{X_i})$ depicts the parent nodes of $\boldsymbol{X_i}$, i.e. the variables $\boldsymbol{X_i}$ is dependent on. The conditional probabilities are represented as parameters $\boldsymbol{\theta_{ijk}} = \boldsymbol{\mathcal{P}}(\boldsymbol{X_i} = \boldsymbol{k} | \mathrm{Pa}(\boldsymbol{X_i}) = \boldsymbol{j}), \boldsymbol{1} \leq \boldsymbol{k} \leq \boldsymbol{s_i}, \boldsymbol{1} \leq \boldsymbol{j} \leq \boldsymbol{p_i}$ with $\boldsymbol{s_i}$ specifying the number of states of $\boldsymbol{X_i}$ and $\boldsymbol{p_i}$ the one of its parent nodes. Given a dataset (set of $\boldsymbol{n} \in \mathbb{N}$ samples) $\boldsymbol{\mathcal{D}} = \{\boldsymbol{D_1}, \ldots, \boldsymbol{D_n}\}$, we are searching the parameters $\boldsymbol{\theta} = \{\boldsymbol{\theta_{ijk}}\}$ of the network's conditional probabilities such that samples taken from the Bayesian network with parameters $\boldsymbol{\theta}$ match the data $\boldsymbol{\mathcal{D}}$ best.

The typical principle is to perform maximum likelihood estimation (MLE), i.e. to search for the parameters $\boldsymbol{\theta}$ which have most likely produced the dataset $\boldsymbol{\mathcal{D}}$. A commonly used learning method, which is based on MLE, is the expectation maximization (EM) algorithm. This approach can also cope with incomplete data and iteratively proceeds an expectation- and a maximisation step until convergence. Starting with an initial estimate $\boldsymbol{\theta^0}$ (e.g. a uniform distribution), the former completes the dataset based on the current estimate whereas the latter re-estimates the parameters $\boldsymbol{\theta^t} \to \boldsymbol{\theta^{t+1}}$ via MLE.

3.3 Inference and strength of influence

After having learned the parameters θ, the Bayesian network offers a compact representation of the conditional probabilities for all influence factors. With the help of the network as a probabilistic reasoning system, the change of each factor's likelihood can be observed if the values of some variables are given. Proceeding inference estimates the updated likelihoods of all dependent characteristics, for instance the quality of the combined modelling in Figure 1. For our consideration of studying the effect that various sub-instances exert on the holistic quality, causal and inter-causal types of inference can be applied.

Inference enables an approach to measure the strength of influence for each node in the Bayesian network based on the conditional probability distributions. In principle, the strength of influence a parent node $X_{par} \in \mathcal{X}$ exerts on a child node $X_{child} \in \mathcal{X}$ is specified by the change in probabilities of X_{child} in case of given evidence for X_{par} compared to the case of no evidence for the parent. In doing so, we assume X_{par} to adopt each possible status individually and measure the particular change in the child's probabilities using a Euclidean distance for probability distributions. Each of these distances is weighted by the likelihood of X_{par} to assume the respective status. The strength of influence a parent node X_{par} has on a child X_{child} thus calculates as

$$I_{X_{par}}(X_{child}) = \frac{1}{\sqrt{2}} \cdot \sum_{j=1}^{s_{par}} \mathcal{P}(X_{par} = j) \sqrt{\sum_{i=1}^{s_{child}} \Big(\big[\mathcal{P}(X_{child} \big| X_{par} = j)\big]_i - [\mathcal{P}(X_{child})]_i\Big)^2} \tag{2}$$

where s_{par} denotes the number of status the factor X_{par} can adapt and s_{child} the number of status for X_{child}, respectively. With $[\mathcal{P}(\cdot)]_i$ indicating the i-th component of a discrete probability distribution, the square root term represents the Euclidean distance of $\mathcal{P}(X_{child})$ and $\mathcal{P}(X_{child} \big| X_{par})$. Furthermore, a normalisation factor in (2) ensures $I_{X_{par}}(X_{child}) \in [0,1]$. Higher values of $I_{X_{par}}(X_{child})$ indicate stronger influences of parent nodes on their children.

4. Practical application

For the evaluation of the presented method, the data of a digital twin of a 3-axis milling operation is used as an example. The generation of these digital twins is described in detail in [18] and provides the data basis for the following method. These digital twins contain all relevant meta and process data acquired during the machining process. The process data includes all measurement and control signals of the machine tool and any additional sensor information that can be determined via additional measurement technology. Specifically, these are the nominal/actual positions, the drive currents, and the nominal/actual speed information of the axes and spindles. All this information is linked to the meta information (tool data, technology data, material data and machine data) [19]. These provide the basis for subsequently integrating technological calculation models (tool engagement models, kinematic models, surface location error models etc.) and thereby generating a digital twin across the complete process chain. In this way, all location- and time-discrete technological information of the tool-workpiece interaction can be calculated and its effect on the produced component can be determined. With the help of Bayesian networks, the factors influencing the position deviations (Δ_x, Δ_y, Δ_z) of the TCP are to be determined as a function of the axis-specific jerk ($\boldsymbol{j_x}$, $\boldsymbol{j_y}, \boldsymbol{j_z}$), the spindle load ($\boldsymbol{L_S}$), the axis-specific drive currents ($\boldsymbol{I_x}$, $\boldsymbol{I_y}, \boldsymbol{I_z}$) and the cutting forces according to the Kienzle cutting force model

$$F_c(\varphi) = k_{c1,1} b h_{(\varphi)}{}^{1-m_c} \tag{3}$$

for the respective milling operations [20]. To determine the cutting forces $\boldsymbol{F_c(\varphi)}$, in addition to the cutting force constants ($\boldsymbol{k_{c1.1}}$,$\boldsymbol{m_c}$), the tool engagement parameters ($\boldsymbol{b}$, $\boldsymbol{h_\varphi}$) must also be determined, which cannot be derived from the pure axis position data of the machine. For this purpose, a material removal simulation (based on a multi-dexel model) based on the actual position of the TCP must be used to determine the time- and location-discrete entry angle $\boldsymbol{\varphi_{in}}$ and exit angle $\boldsymbol{\varphi_{out}}$ together with the chip width of the undeformed chip $\boldsymbol{b}$. For the subsequent calculation of the mean undeformed chip thickness $\boldsymbol{h_m}$ from

$$h_m = \frac{1}{\varphi_c} \cdot \int_{\varphi_{in}}^{\varphi_{out}} h_{(\varphi)} d\varphi = \frac{1}{\varphi_c} \cdot f_z \cdot \sin\kappa\, (\cos\varphi_{in} - \cos\varphi_{out}), \tag{4}$$

the tooth feed $\boldsymbol{f_z}$ must also be calculated. Using the time stamp of the individual tool positions, the feed per tooth $\boldsymbol{f_z}$, can be calculated by numerically differentiating the actual positions in $\boldsymbol{x}$-, $\boldsymbol{y}$- and $\boldsymbol{z}$-directions and using the corresponding actual spindle speed. The calculation of the mean cutting forces $\boldsymbol{\overline{F}_c}$ is specified in [21].

As test machine, a WEMAS VZ-1250 Quick with a HEIDENHAIN TNC620 control was selected. An aerospace component was manufactured on this test machine, for which four successive manufacturing steps are described here as examples (see Figure 2). For all manufacturing steps, the relevant process data is measured with a sampling rate of 2 ms. The exact process for acquiring process data is described in [18]. In the first process step (*face milling operation – T25*), the raw material is face milled, for which the feed movement mainly takes place in the x-direction. In the following processing step (*drilling operation – T26*), a drilling operation takes place, where the main feed movement is in the z-direction. In the subsequent contour milling (*contour milling operation - T27*) and finishing (*end milling operation – T28*), no main feed movement direction can be identified, which is why the determination of the factors influencing the position deviations is particularly relevant for these machining operations.

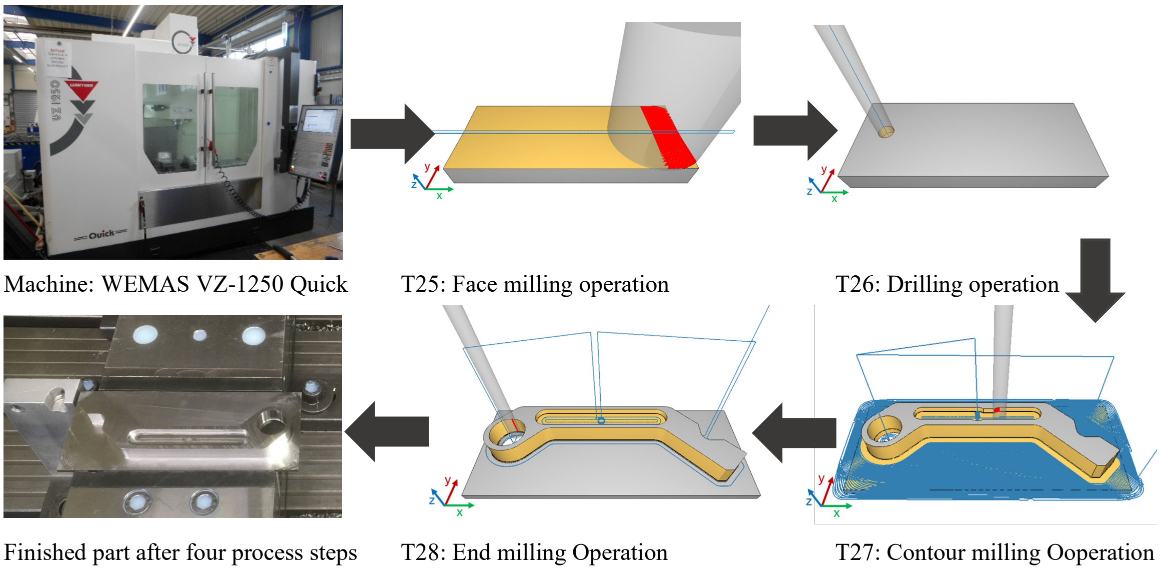


Figure 2: Application scenario using the example of a milling operation

4.1 Network structure and data preparation

Based on expert knowledge, the structure of the Bayesian network is constructed of three layers. The upper layer contains the jerk values of the x-, y- and z-axis and the mean cutting force. It is intended to evaluate the influence of these values on the total positional deviation and the axial deviations. Since preliminary computations show the axial deviations are influential dominated mainly by the jerk on the same axis, these axial deviations are equipped with only a single parent node in order to simplify the network structure. Differences in the strength of influence with respect to the axial jerks are considered on the total deviation. The same procedure applies for the positional deviations in the second layer whose influence on the nominal amperages in the bottom layer are examined. To also include a more global approach of estimating influence, the spindle load is included into the bottom layer and is seen as a consequence of the positional axial deviations. Figure 3 summarises the structure of the constructed Bayesian network. Alongside, other network structures may be considered and compared against each other.

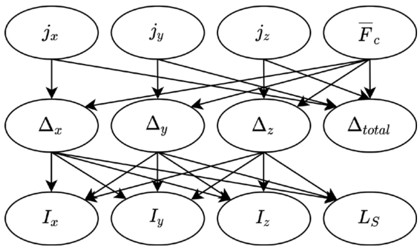


Figure 3: Structure of the Bayesian network

For each part process the recorded data of the presented factors are assembled in a csv-file. A record of factor values belonging to a certain instant of time produce a sample for the parameter learning, i.e. we use each instant of time as an observation. Due to all the factors are on hand as continuous variables, a pre-processing is required to discretise the data. We subdivide the individual factors into three classes. The smallest 10% of values are classified as »small« whereas the largest 10% of values are considered as »large«. The remaining part is squeezed in the group »medium«. This choice is motivated on the one hand to ensure that each combination of the variables' discrete values holds sufficient data records and on the other hand to depict the effects of lighting particular small respectively large measurements of certain influential variables. The nominal amperage streams of data ι_{nom} are primarily smoothed via a running average and centred corresponding to

$$\iota = \left| \iota_{nom} - \iota_{nom}^{smoothed} \right| \tag{5}$$

where furthermore only the absolute amperage values are considered due to the applied discretisation criterion. The latter also deploys on the jerk values. The cutting force is sampled in lower frequency than other measures. Hence, the interim values are interpolated using cubic C^2-splines.

4.2 Learning process and influence evaluation

To handle the Bayesian networks, we use the R-package »bnlearn« as well as the software »GeNIe«. The parameters of the network are learned with the EM-algorithm, which in our case corresponds to the MLE-method using an initial estimate since the considered machining process provides a complete and extensive dataset. Having learned the parameters in the form of the conditional probability tables, we examine the strengths of influence for each node (parent-child-relation) as described in section 3. Figure 4 illustrates the strength of influence calculation by considering changes in the child's probabilities, if evidence on the parent is added, and the comparison of multiple influential variables on a common child node.

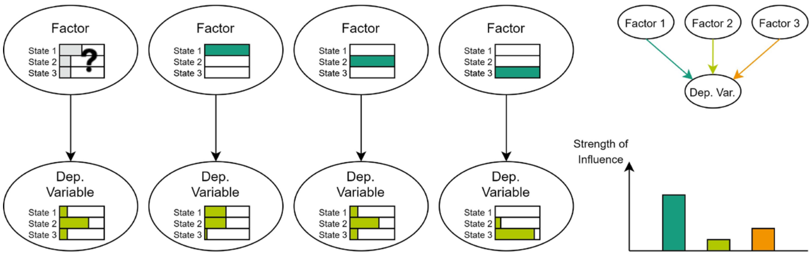


Figure 4: Visualisation of the influence evaluation using (2) and comparison for multiple influencing factors of a common depended variable to identify the dominating factor

An approach to present such single-edge influences combined with the network structure is to watch the strengths of influence as weights to the edges. Thus, we get a global comparison of all degrees of influence and at the same time as local contrasts in relation to a certain node and its parents. To illustrate how to evaluate the dominating influence factors of a specific measure, we exemplary consider the effects of the axial jerks and the cutting force on the total deviation and furthermore the influence of the axial deviations on the spindle load.

4.3 Process characteristics and results of influence analysis

The holistic application is analysed by a group of experts and the characteristics of each single operation (T25 - T28) are exposed for a comparison with and justification of the findings of the presented approach. The face milling operation (T25) is characterised by uniform motion, which leads to small differences in the influence strengths of axial jerks. Slight influences of the jerk of the z-axis result from the approach movements, whereby the approach movement was realised at high speeds (rapid traverse). The axis determining the feed direction compensates for the main influence on the workload of the spindle. The face milling operation exhibits the x-axis whereas the drilling operation (T26) shows the z-axis as the determining one. Furthermore, T26 does not show noteworthy changes in velocity. Together with the just punctual processing of the drilling operation, this results in few influence of the axial jerks on the total deviation. In contrast, we see in the contour milling operation (T27) and end milling operation (T28) comparatively rapid changes in machining velocity and direction, which strengthen the influence of the cutting force. Although the end milling operation (T28) holds smaller forces compared to the contour milling operation (T27), but particularly high axis speeds (finishing operation). Both the contour and end milling process exhibit the x-y-plane as the cutting plane that therefore provides the feed force direction. Hence, the deviations of the corresponding axes should represent the most important influence factors. The same applies for the nominal amperages.

The appertaining strengths of influences for the factors deviation, calculated using (2) with the presented approach, are depicted in Figures 5 and 6.

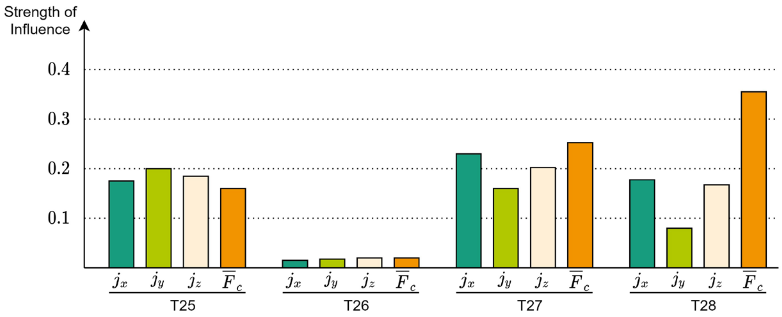


Figure 6: Strength of influence evaluation for axial jerks and Kienzle cutting force on the total deviation Δ_{total} for each process step individually

Figure 5 predicates few differences in the influence strengths for the face milling operation (T25) and drilling operation (T26), which confirms the smooth and regular movement characteristics of both processes. Due to the small-area processing in the drilling operation (T26), the influence strengths of the factors in this operation are particularly small. In contrast, we recognise a significant influence of the cutting force during the contour milling operation (T27) and end milling operation (T28) due to permanent changes in machining velocity and direction. Because of these characteristics, the comparison of the jerks is especially interesting. Thereby, the influence of the jerks of the x- and z-axis dominate the jerk of the y-axis. It has to be considered that the z-direction is considerably influenced by rapid traverse return strokes which do not influence the surface quality. Therefore, we conclude for this application, the greater the speeds, the more dominant the influence of the process-related factors (i.e. the cutting forces) are.

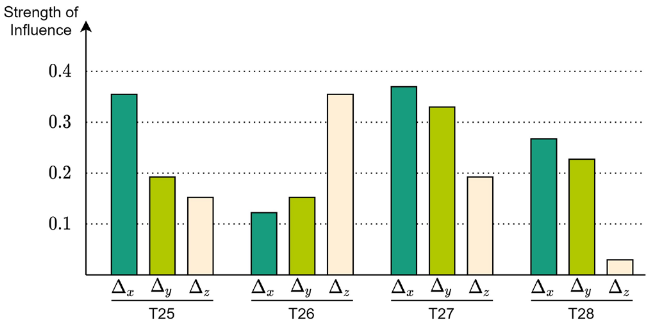


Figure 5: Strength of influence evaluation for axial deviations on the workload of the spindle L_S

Considering Figure 6, the face milling operation (T25) and drilling operation (T26) reflect that the highest estimated influence strengths on the spindle load belong to the axes corresponding to the feed direction. For contour milling operation (T27) and end milling operation (T28) we can identify the deviations of the x- and y-axis as the mainly impacting factors of the workload. With regard to the nominal amperages, the deviations of the x-axis affect the amperages most for the face milling operation (T25), whereas the contour milling operation (T27) and end milling operation (T28) point out the x- and y-axes as the dominating influencing factors, which covers the assessments by the group of experts.

5. Quantum computing and Bayesian networks

Both forward and inverse inference require a high computational effort in Bayesian networks including a large number of nodes. A method to handle such large-scale networks as well as to consider multiple structuring approaches in parallel consists in the technology of quantum computing, which is based on the usage of quantum bits (so called »qubits«). One decisive principle of qubits is superposition, i.e. the capability of qubits to be in the »classical« states of a bit 0 and 1 simultaneously. This characteristic forms the basis to represent the nodes together with their parameters as qubits. The superposition states of a qubit $|\boldsymbol{\psi}\rangle$ are denoted as $|\boldsymbol{\psi}\rangle = \boldsymbol{\alpha}|\mathbf{0}\rangle + \boldsymbol{\beta}|\mathbf{1}\rangle$, where in quantum computing the classical basis states 0 and 1 are represented in Dirac notation. The complex coefficients $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ specify the probabilities of the qubits to segue into a basis state when being measured with $\mathcal{P}(|\psi\rangle = |0\rangle) = |\alpha|^2$ and $\mathcal{P}(|\psi\rangle = |1\rangle) = |\beta|^2$ at what the coefficients of a superposition have to hold $|\alpha|^2 = |\beta|^2 = 1$.

To model a Bayesian network as a quantum circuit, each node is assigned to one or multiple qubits, depending on the number of discrete states the corresponding variable can be in. More precisely, $\lceil \log_2(n) \rceil$ qubits are required to map the n different discrete states of a random variable. The encoding of the probabilities to the parameters is done via rotational gates $R_y(\Phi)$ (for nodes that are not dependent on other factors) and m-times controlled rotational gates $C^m R_y(\Phi)$ [22]. The rotation angles Φ are determined by the parameters of the respective node in the Bayesian network. The second important principle of quantum computing, entanglement, enables to set the dependencies between multiple factors. The corresponding qubits are entangled via the controlled rotational gates, where m corresponds to the number of parent nodes. The assembling of the Bayesian network is sometimes also referred to as qsample encoding [23].

Since in the presented milling process each node holds three states, two qubit are required to encode a factor. Once the network is encoded, we can set evidence via resetting of qubits (analogous to the procedure of Figure 4) and entangle them with ancillary qubits to encode the changed probability distribution of the child node. Then, the change in probabilities is measured via a swap test, which is a procedure of determining how much the states of two quantum registers differ. As the probability distributions are now directly encoded in particular quantum states, the swap test delivers the strengths of influence estimations via (2). The swap test provides a measurement parameter that is positively correlated to the actual distance. Such parameters are calculated as the inner product between normalised vectors, which consist in the qubits' superposition states.

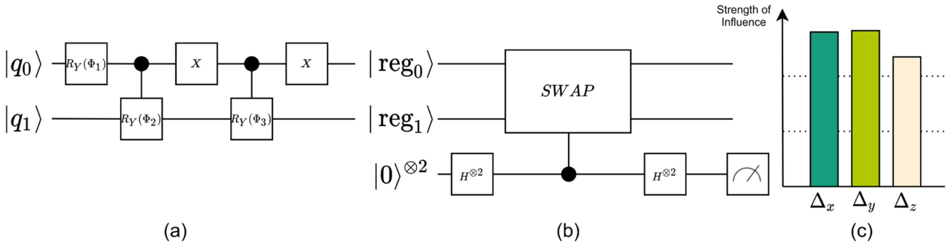


Figure 7: Parameter encoding for a node via rotation angles (a) and measurement of changes in probabilities via swap test (b) as two essential blocks of the quantum circuit, the results for influence estimation on spindle load for T27 (c)

In Figure 7, the parameter encoding in two qubits is shown for a 3-state node. Furthermore, the adequate swap test takes two quantum registers of two qubits each and delivers the result by measuring an ancillary register of corresponding size. To demonstrate the merely capability of the presented quantum algorithm, we implemented this approach on the IBM quantum systems including a quantum simulator with low error rates as well as real quantum hardware. We exemplary evaluated the setup of Figure 6 on the spindle load by the quantum simulator. As an extract, the results for the contour milling (T27) are included in Figure 7, also showing the x- and y-axis as the dominating influence factors.

Besides the advantage in storage, quantum computing offers a potentially quadratically improved runtime performance over conventional computing methods when it comes to inference, which is a subroutine to estimate the strengths of influence. This advantage is gained using a quantum version of rejection sampling as an approximative inference algorithm. The quantum method and its quadratic speedup is thereby based on the technique of amplitude amplification: Starting with a superposition of all possible solutions, the probability amplitude of the correct solution is reinforced gradually with a simultaneous decrease of the other amplitudes. At the end of this procedure, the true solution is given with very high probability.

6. Conclusion and Outlook

This article described how the characteristics of machining processes can be represented as Bayesian networks. After having formalised the measurement to evaluate strengths of influence, we demonstrated the applicability of this approach using a digital twin of a milling operation. In order to utilise available real data of a machining process, we focussed the practical application on a single digital twin, but the presented methodology is directly transmissible to the context of coupled digital twins. The results of the presented approach reflected, on the one hand, the findings on the characteristics of the considered machining process, which underlines the applicability in a correct manner. On the other hand, additional evidence like an increased influence of the process-related factors at higher speeds was obtained. Moreover, the functionality of Bayesian networks and influence estimation is integrated in the context of quantum computing and the consequent potential advantages are depicted. Our current and future research directions are to develop an application to a holistic process chain and domains of different interrelated manufacturing processes along the value chain, and to develop broader quantum circuits capable of covering more complex application scenarios by combining Bayesian network construction and influence estimation in a single quantum algorithm.

References

[1] Neugebauer, R., Ihlenfeldt, S., Frieß, U., Wabner, M., Rauh, S., 2012. New high-speed machine tool structure by holistic mechatronic systems design. Procedia CIRP (1), 307–312.

[2] Fenton, N., Neil, M., 2018. Risk Assessment and Decision Analysis with Bayesian Networks. CRC Press, Boca Raton, Florida.

[3] Jordan, M., Kleinberg, J., Schölkopf, B., Jensen, F.V., Nielsen, T.D., 2007. Bayesian Networks and Decision Graphs. Springer New York, New York, NY.

[4] McLachlan, S., Dube, K., Hitman, G.A., Fenton, N.E., Kyrimi, E., 2020. Bayesian networks in healthcare: Distribution by medical condition. Artificial intelligence in medicine 107, 101912.

[5] Greppi, A., Giuli, M.E. de, Tarantola, C., Montagna, D.M., 2018. Bayesian Networks for Financial Market Signals Detection. In: Mola F., Conversano C., Vichi M. (eds) Classification, (Big) Data Analysis and Statistical Learning. Studies in Classification, Data Analysis, and Knowledge Organization. Springer, Cham., 219–226.

[6] Hosseini, S., Ivanov, D., 2020. Bayesian networks for supply chain risk, resilience and ripple effect analysis: A literature review. Expert systems with applications 161, 113649.

[7] Jones, B., Jenkinson, I., Yang, Z., Wang, J., 2010. The use of Bayesian network modelling for maintenance planning in a manufacturing industry. Reliability Engineering & System Safety 95 (3), 267–277.

[8] Carbery, C.M., Woods, R., Marshall, A.H., 2018. A Bayesian network based learning system for modelling faults in large-scale manufacturing. Proceedings - 2018 IEEE International Conference on Industrial Technology, 1357–1362.

[9] Sun, Y., Qin, W., Zhuang, Z., 2020. Quality consistency analysis for complex assembly process based on Bayesian networks. Procedia Manufacturing 51 (7648), 577–583.

[10] Nannapaneni, S., Mahadevan, S., Rachuri, S., 2016. Performance evaluation of a manufacturing process under uncertainty using Bayesian networks. Journal of Cleaner Production 113 (3), 947–959.

[11] Sardeshmukh, A., Reddy, S., Gautham, B.P., 2019. Bayesian Framework for Inverse Inference in Manufacturing Process Chains. Integr Mater Manuf Innov 8 (2), 95–106.

[12] Klein, A., 2020. Efficient Bayesian Hyperparameter Optimization Dissertation at Freiburg University.

[13] Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., Deng, S.-H., 2019. Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. Journal of Electronic Science and Technology 17 (1), 26–40.

[14] Bustillo, A., Pimenov, D.Y., Mia, M., Kapłonek, W., 2021. Machine-learning for automatic prediction of flatness deviation considering the wear of the face mill teeth. Journal of Intelligent Manufacturing 32 (3), 895–912.

[15] Yahmadi, R., Brik, K., Ammar, F., 2018. Causal tree analysis for quality control of the lead acid battery manufacturing process. International Journal of Energy Research 42 (4), 1738–1759.

[16] Deresse, N.C., Deshpande, V., Taifa, I.W.R., 2020. Experimental investigation of the effects of process parameters on material removal rate using Taguchi method in external cylindrical grinding operation. Engineering Science and Technology, an International Journal 23 (2), 405–420.

[17] Scutari, M., Graafland, C.E., Gutiérrez, J.M., 2019. Who learns better Bayesian network structures: Accuracy and speed of structure learning algorithms. International Journal of Approximate Reasoning 115 (3), 235–253.

[18] Hänel, A., Schnellhardt, T., Wenkler, E., Nestler, A., Brosius, A., Corinth, C., Fay, A., Ihlenfeldt, S., 2020. The development of a digital twin for machining processes for the application in aerospace industry. Procedia CIRP 93, 1399–1404.

[19] Caesar, B., Hanel, A., Wenkler, E., Corinth, C., Ihlenfeldt, S., Fay, A., 2020. Information Model of a Digital Process Twin for Machining Processes. 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), 1765–1772.

[20] Kienzle, O., Victor, H., 1957. Spezifische Schnittkräfte bei der Metallbearbeitung. In: Werkstofftechnik und Maschinenbau 47 (H5), 224–225.

[21] Hanel, A., Wenkler, E., Schnellhardt, T., Corinth, C., Brosius, A., Fay, A., Nestler, A., 2019. Development of a Method to Determine Cutting Forces Based On Planning and Process Data as Conribution for the Creation of Digital Process Twins. MM SJ 2019 (04), 3148–3155.

[22] Borujeni, S.E., Nguyen, N.H., Nannapaneni, S., Behrman, E.C., Steck, J.E., 2020. Experimental evaluation of quantum Bayesian networks on IBM QX hardware. 2020 IEEE International Conference on Quantum Computing and Engineering (QCE), 372–378.

[23] Schuld, M., Petruccione, F., Schuld, M., Petruccione, F., 2018. Supervised Learning with Quantum Computers. Springer International Publishing, Cham.

Biography

Maximilian Selch, M.Sc. (*1995) completed his studies in mathematics at Dresden University of Technology in 2020 and is a research associate at the Fraunhofer Institute of Machine Tools and Forming Technology IWU.

Albrecht Hänel, Dipl.-Ing. (*1988) was born in Dresden in 1988 and graduated in 2014 with a diploma in production engineering at Dresden University of Technology. After completing his studies, he worked as a research associate at the Institute of Manufacturing Science and Engineering at Dresden University of Technology until 2019. In 2019, he moved to the department of Cyber-Physical Production Systems at the Fraunhofer Institute for Machine Tools and Forming Technology IWU.

Uwe Frieß, Dr.-Ing. (*1981) studied machine tools an applied mechanics at the Chemnitz University of Technology until 2007 and worked in the industry thereafter until 2009 when he returned to the University. He finished his Ph. D. Thesis in 2020 on fuzzy clustering of machine states for condition monitoring and is currently group leader for Data analytics and hybrid State models at the Fraunhofer Institute of Machine Tools and Forming Technology IWU.

Steffen Ihlenfeldt, Prof. Dr.-Ing. (*1971) is professor at the chair of Machine Tools Development and Adaptive Controls at Dresden University of Technology since 2015. After studying mechanical engineering at Braunschweig University of Technology, Steffen Ihlenfeldt worked as head of the department Cyber-Physical Production Systems at the Fraunhofer Institute of Machine Tools and Forming Technology IWU and is institute director of the Fraunhofer IWU since 2021.

2nd Conference on Production Systems and Logistics

Potentials Of Blockchain In Crowdsourcing Platforms – An Outlook For Industrial Services

Nick Große[1], Martin Eisenmann[2]

[1]*Chair of Enterprise Logistics, Dortmund, Germany*
[2]*Social Research Center, Dortmund, Germany*

Abstract

Companies increasingly outsource services with the intention to disseminate risks and workload (problem) with other organizations. Reasons for that may be the lack of internal expertise, reduced execution costs, and network effects such as focus on the core business. Crowdsourcing is a way of disseminate the workload, utilizing external expertise and solving problems of a project with other, partly unknown, network participants. The goal of crowdsourcing is to separate responsibility and to balance the workload of employees (peer) or to make use of suitable external workforces coordinated by network mechanisms (platform). Crowdsourcing appears in an ambivalent way and needs regulating and participatory structures of employment. Due to the fact that a failure of a single entity may lead to the failure of the whole project, the mutually unknown participants have to rely on each other's quality (performance). Cooperations are prone to information asymmetry and its corresponding uncertainty in terms of the partners' behaviour which leads to the question of trust between the cooperating partners (principal and peer). This paper addresses the entities of a crowdsourcing system under the scope of the principal agent theory and its underlying behavioural assumptions. Five essential elements will be derived: principal, peer, problem, platform and performance (5 Ps). Based on this, the potentials of the blockchain technology will be explored by reflecting its functionalities to the derived elements and its contributions to ensure trust despite of information asymmetry in crowdsourcing platforms.

Keywords

Crowdworking; Regulation; Behavioural Uncertainty; Industrial Services; Blockchain; Trust

1. Introduction

In the course of the "fourth industrial revolution" [1] or the "Second machine age" [2], companies in industrial sectors must adapt to their changing environments. One of the most challenging changes in the digitalized industry and in the development of smart maintenance [3] is the growing complexity of machines and plants (assets) as well as the intensive usage of data and new technologies [4, 5]. Despite this complexity, industrial services and the in-house maintenance departments have to guarantee the reliability and availability of all assets which are required for the production [3]. The growth of complexity combined with the increasing demand of skilled workers [6] requires adaptability also for industrial services and in-house maintenance departments [7]. These developments and changes, in addition to rationalization efforts, are causing companies to reduce complexity, e.g. through intelligent technology and data usage, and through the recruitment of external personnel which is appropriately qualified. As an alternative to the "onboarding" of highly qualified skilled workers, the widely increasing use of crowdsourcing platforms is an option [8]. On those platforms, tasks and jobs are offered (or outsourced) to an undefined mass of potential (crowd-)workers which indicates the change in work organization [9].

Within those economic relations or networks there are lots of uncertainties and the risk of a lack of trust. Against the background of the principal-agent theory (P.A. theory), the aim of this paper is, firstly, to analyze and redefine the core elements that characterize the interaction of the network around the phenomenon of crowdworking and, secondly, to answer the question of how the trust deficits can be overcome by using the potentials of the "Blockchain Technology" (BCT) in the case of crowdworking platforms. This contribution begins by analyzing the theoretical background which consist of the outline of the P.A. theory and its core-elements as well as the definition of relevant and necessary terms which are essential for a general understanding (**chapter 2**). Subsequently, we analyze the core elements of network-like crowdworking platforms and elaborate which elements should be added in the sense of the P.A. theory as well as how crowdworking platforms can be regulated by participatory mechanisms. For that purpose, we summarize the elements as the 5 Ps of a crowdworking ecosystem and focus on the option of implementing such an ecosystem in the economic segment of industrial services (**chapter 3**). In the next section, we work out the potential of BCT and how it can mitigate or eliminate behavioural uncertainties in the case of industrial services (**chapter 4**). This is followed by a conclusion of the paper which consists of a short summary of our findings and an outlook (**chapter 5**).

2. Theoretical background

The following chapter deals with the context specification and the underlying theoretical basis of the paper. This chapter briefly introduces the terms network and crowd, and explains its linkages to the aspects of trust and information asymmetry in such environments. The chapter is concluded by reviewing recent works about past research dealing with the consideration of blockchain and crowdsourcing.

Crowdsourcing can be described as an outsourcing activity and is a composition of the words 'Crowd' and 'Outsourcing' and implies the outsourcing of a task or a plenty of human workers (Crowd) [10,11]. The main idea behind crowdsourcing is to exploit the potential of knowledge and competence of internal and external employees [11]. The coordination structure of the crowdsourcing settles in the hybrid form between market and hierarchy [11] can be compared to the hybrid form called networks [12,13]. Networks in turn are basically social networks or arrangements and can be differentiated based on their kind of orientations such as knowledge networks, intercompany networks etc. [14]. For that reason, terms such as intercompany networks and crowd will be used synonymously in this paper.

The success of a cooperation in such an environment depends on the underlying coordination effort which can be quantified through transaction costs [15,16]. The amount of transaction costs is affected by the information asymmetry [17]. Managing such crowdsourcing arrangements poses challenges in terms of uncertainties among the network participants due information asymmetry which can be described via the PA theory [18,9]. An increased uncertainty about the participants' behaviour mitigates trust in network cooperations [16]. On an interorganizational level, trust consists of personal and systemic trust (organization) [19]. Due to that, the emergence and maintenance of trust in such an environment is essential for preventing the occurrence of negative-scenarios such as adverse selection [20] which is characterised by a total displacement of honest participants and hidden characteristics [21]. This scenario is also an indicator for instability of such an arrangement [17]. For that reason, infrastructures such as the BCT receive closer consideration in terms of information asymmetries in interorganizational arrangements [22,23]. Blockchain, a subset of Distributed Ledger Technologies (DLT), can be described as a "*distributed database that is practically immutable by being maintained by a decentralized P2P network using a consensus mechanism, cryptography and back-referencing blocks to order and validate transactions*" [24, p.13].

Past research investigated the potentials of blockchain in crowdsourcing environments such as in the fields of energy systems for finding an equilibrium in the energy allocation [25]. Another example is the application on reputation systems to ensure trust within crowdsourcing platforms [26]. Further contributions are coping with challenges in terms of user-generated content, but without reference to blockchain [27]. In the fields of industrial services, there are attempts in which crowdsourcing is considered in the domain of maintenance scenarios. However, only a few studies dealt with blockchain-based applications in the fields of the maintenance [28], even less in combination with crowdsourcing such as *DING ET AL*. [29]. The contributions from those studies serve as a basis and motivation for considering them in regard to the themes addressed in this paper.

3. Crowdsourcing ecosystems and regulating participation mechanisms

The following section of the paper provides definitions of crowdsourcing and outlines the crowdsourcing environment. With regard to the organization of work, we also focus on the regulating mechanisms which can enhance the participation of crowdworkers. Additionally, we list up the core elements of a crowdsourcing ecosystem and expand those elements by adding two more elements which describe the system of crowdsourcing more precisely. The basis for this is the above mentioned PA theory.

3.1 Definition and environment of crowdsourcing ecosystem

The term "crowdsourcing" appeared in 2006 for the first time and is associated with *JEFF HOWE* who used it in the Wired Magazine [10]. In a first alignment of the term, crowdsourcing is defined in our contribution as follows: "Crowdsourcing refers to the outsourcing of certain tasks by a company or generally an institution to an undefined mass of people by an open call, which is mostly made via the Internet" [9, p. 87]. The basic technology is of course the internet as well as information and communication technologies which connects the different elements with each other [30].

For the principal or contractor, the concept of crowdsourcing subsequently leads to lower transaction costs for receiving information and knowledge and thus to flexible, fast and favourable employment options. Hence, hierarchy related coordination mechanisms in an organization are therefore replaced by a market-based form of coordination [31, 9]. Examples for crowdsourcing platforms are: InnoCentive, NineSigma, Amazon Mechanical Turk, Microworkers, Crowdworx, ChaCha, Atizo etc., which cover several parts of the value chain (e.g. computer science, research and development, finance, marketing, engineering) [9, 30]. It has to be noted that there is a difference between internal and external crowdsourcing. The former focuses on internal employment which uses an intra-organizational and often global platform, whereas the latter crowdsourcing form consists of employees external to the company [30, 9]. For this article, especially the external crowdsourcing is relevant. Furthermore, crowdsourcing can be differentiated into a cooperative-based approach and a competition-based (e.g. 99 designs) approach which can be further differentiated into an outcome- or time-oriented form of crowdsourcing [9]. Typically, the crowdsourcing process takes place in the following phases: concretisation of tasks (1), selection of the crowdworker (2), task performance (3), aggregation and solution (4), payment (5) [9, 31].

With regard to the domain of industrial services or field services, we found only a few platforms which coordinate the interests of principals and contractors which shows that crowdsourcing in this economic sector is new but growing. In the case of field services, the app of the eponymous company "mila" is worth mentioning [32]. This platform-like network brings private clients, who order different solution for technical problems as well as small and medium sized or even big companies (e. g. Vodafone, Swisscom, Conrad etc.), together with private contractors or professional service technicians. Depending on the problem, the professional technician works remote or on-site. The example of the "mila" platform operates in combination with a field service management app which coordinates the data and information about order, location, problem, payment etc. Another example for platform- or market-based coordination of principals and contractors in the domain of industrial maintenance services is the software-based platform for field service management called “crowd service”. Registered customers can book certified and highly qualified service technicians and hire them using a field service management app which connects the freelance service technician with the customer [33]. The booked technician scans the QR-code of the broken asset or machine which has to be maintained or repaired, obtaining the relevant data-set via app [34]. Especially for the principal, the usage of crowdsourcing platforms brings some advantages. On the one hand, the following aspects are beneficial to the principal: access to (external) know-how, innovative solutions, high process velocity, decrease of costs, increase of flexibility, customer satisfaction etc. On the other hand, there are some disadvantages such as: high effort of defining tasks, in-transparency of the cost structure, loss of control regarding the activity, knowledge drain etc. [9, 30].

3.2 Regulating crowdsourcing as a new form of work organization

“*Crowdsourcing is not merely an innovative concept for the distribution and execution of corporate tasks, but rather a completely new way of organizing work, which can be accompanied by (sometimes radical) changes on both the employee and employer side*” [9, p.11]. The more market-based coordinated than hierarchy-based crowdsourcing networks implicate huge changes for work and the organization of work which can be differentiated in beneficial and conflictual changes [31]. Beneficial changes for crowdsources are: new employment opportunities, autonomous work, peer exchange and greater flexibility, whereas conflictual changes occur in the following forms: low payment, monotonous and standardized activities, increased potential for control and lack of legal framework (working) conditions [31].

Crowdsourcing obtains its power from the knowledge and decentral-located resources of various independent and parallel working people which enable fragmentation, standardization and automation of tasks [31]. Unspecified working time, employee participation rights and qualification opportunities lead to further questions of better working conditions and how they can be enabled or regulated. Union representatives have been interested in issues related to crowdsourcing for some time [35, 30, 36] and of how “good work” should be designed in digitalized workspaces [35]. The critic of unions focuses on unfavourable working conditions and the unregulated forms of earning. The usage of (unknown) external workforces has enormous consequences for the internal and often union regulated workforces (e. g. decreasing wages or “precarity” [30]). “*In the context of crowdsourcing initiatives, work activities are removed from a legally regulated, company employment relationships and shifted into a kind of legal vacuum*” [30, p. 3]. Work and employment on platforms do neither take place in institutions which are regulated by labour- or collective bargaining law nor do they provide for participation rights [30].

Another question that unions try to answer is how much responsibility platforms take to implement fair working conditions, because the responsibility diffuses into the (digital) market and does not lie in the hands of employer associations and unions anymore, but to the crowdsourcing platform. To avoid the establishment of so called “sweatshops”, unions search the dialogue with crowdsourcing platforms, adapt labour and social law and try to articulate a new employment and company definition [35; 36]. For employee representatives, crowdsourcing as a digitalized form of outsourcing means: competitive pressure, competition for locations, disputes with relocation and often a lowering of working standards in terms of pay or working hours in order to save jobs [35]. Regarding the network of principals, platforms and crowdsourcing peers, the challenge is to bridge the existing gap of information asymmetries between those stakeholders which could lead to, on the one hand, fairer working conditions for employees and, on the other hand, to lower risks and more transparency for the principal. The platforms fulfil a key role in this network, because they are coordinating and controlling the affairs of the other two stakeholders. This leads to the question of how to unlock positive aspects and creative potential of the crowdsourcing phenomenon, while at the same time fair conditions of digital work needs to be shaped and designed [35].

3.3 5 P´s of a crowdsourcing ecosystem in industrial services

Against the background of the P.A. theory, *LEIMEISTER AND ZOGAJ* [9, p. 18-21] collect various definitions of crowdsourcing and they work out core elements which are observable within this network-like interaction. The network consists of: the crowdsourcer offering the order who we define as the *principal*; the crowdsourcee executing the order for payment who we define as the *peer*; the (external) crowdsourcing *platform* which is an intermediary institution organizing and coordinating orders and fitting activities as well as the process which contains the transaction, communication and execution of the formulated *problem* and its solution [9]. With regard to the underlying research question of this article, we assume that within the focused network of a crowdsourcing ecosystem there is one more element which could have an influence on the network. Focusing the economic sector of industrial services, those elements have to be redefined and must be supplemented by one element: the *performance*. Hence, we will define all five elements in an enhanced concept which we call the 5 Ps of crowdsourced industrial service. In the following paragraphs, the five elements are described and illustrated in Figure 1:

Platform: The connecting structural element between principals and peers is the (external) IT platform which shifts the hierarchy-based to a market-based coordination of labor [9]. The platform processes, coordinates and controls the interactions. Platforms can either be established by the principal itself or by an external provider to fulfil the required tasks and projects as a service. Such an external provider is also called an "intermediary" [30]. In contrast to internal crowdsourcing platforms, external platforms do not guarantee employee rights such as: regulating termination, vacation, minimum wage, social insurance etc. [35] which poses challenge in the next years for platforms, principals and peers. For platforms which provide industrial or field services we found only a few existing providers, yet. The above mentioned platforms "mila" and "Crowd Service" [32; 34] are first platforms which use a field service management software or app to connect the stakeholders with each other. This allows a platform-based allocation of the principal's needs (e.g. fixing a broken machine) with the required service technician, mechanic or mechatronic engineer whose appropriate qualifications and profile fit to the described problem [34]. Platforms pave the way to a pool of favourable resources such as: know-how and abilities (of service technicians) and mitigate the risks of communication between the peers and the principal [9].

Principal: The crowdsourcer or the principal initiates the process by defining the problem to be solved and by specifying the tasks to be processed. Furthermore, the principal determines incentives and exploits possible solutions while the peers select and process the tasks provided. The principal formulates an order and communicates it to the platform. After receiving the solution, the principal has to verify the right quantity and quality of the outcome, controls and evaluates the result of the peer [9]. The strategy of the principals is to use external workforces via internet or cloud services to decouple the tasks and knowledge and to make use of the external resources which are independent from time, place and organizational boundaries [30].

Peer: The crowd is defined as an undetermined amount of people which we call peer. The peer group could probably consist of all potential internet users, but those are in most cases specific and often certified peers. Hence, the size and structure of the peer group is determined by the call which could be completely open or restricted. In the case of a restricted call a specific group of peers is addressed who is "*characterized, for example, by certain qualifications or personal characteristics*" [30, p.9]. In the focused economic sector of industrial or field services, the group of peers consists of a "*pool of skilled, freelance service technicians"*, mechanics, mechatronics engineers, who are *"available to respond to service calls*" by using their necessary know-how and abilities to handle various problems [34, p. 7].

Problem (and solution): The problem and the definition or the concrete framework of the required tasks to be worked on determines the incentive structures and evaluates the solutions [9]. The principal defines the problem and peers generate a concrete and customized solution for it. For the quantitative definition of the right output or solution (e.g. scheduled working time to repair a broken machine) there are often so called

"Service-Level Agreements" or "Service specifications" which fix time, used material, wages, tools, needed certifications or rights of access [37].

Performance: The performance is the fifth core element we found to be established in a redefined concept of crowdsourcing regarding the economic sector of industrial and field services. In contrast to the usage of only quantitative elements like the "Service-Level-Agreements", qualitative elements of a service technician have to be taken into account as well. Those are for example: the ability of guaranteeing the customer's satisfaction, process or market orientation [37]. Other factors which are of great importance for a good performance are know-how, work experience, qualifications and certified skills as well as a sense for business which are requirements of a well performing service technician [26]. Therefore self-marketing and flexibility become more and more meaningful as well as the profiling of required skills, know-how and experience, qualifications etc. on corresponding platforms like the above mentioned. The skill set of the freelance service technicians must be transparent to get picked by a principal [37, p. 10].

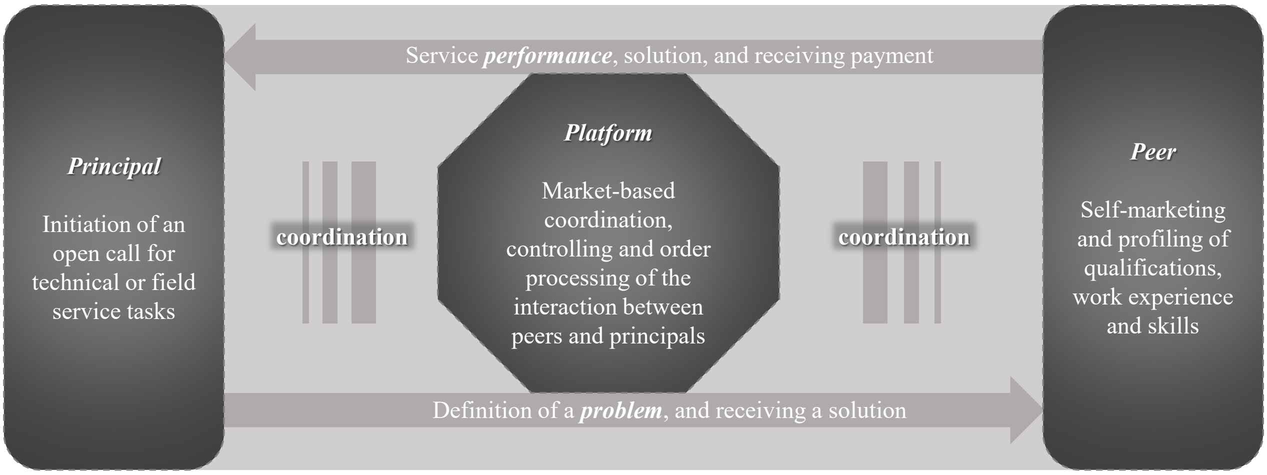


Figure 1: 5 P s of a crowdsouring ecosystem: Focusing the industrial service; own illustration; in accordance with: [9, p.17-21]

4. Potentials of BCT for crowdsourcing in industrial services

In the following chapter, the projection of the derived system features of crowdsourcing systems with regard to the BCT adoption receives closer consideration. The chapter consists of three subsections. First, based on recent literature, a short introduction into the trust-building components of a blockchain is introduced, followed by an alignment of the system features into the correlations between the system components. To propose how the trust-inherent functionalities of BCT can be considered and applied within a specific application field, the industrial service will be used as an example.

4.1 Short introduction into the trust-building properties of a blockchain

Past research repeatedly showed that a blockchain consists of trust-building characteristics which are analyzed by *Wieninger* [23] who states that blockchains are on a technological layer *consensus mechanism*, *decentralization* and *cryptography*, whose combinations can lead to the systemic characteristics *transparency* and *immutability*, serving as a basis for application-based trust-characteristics described by *data-integrity* and *process-integrity* (cf. Figure 2) [23]. *Decentralization* indicates that no central authority is needed for the network organization [38]. *Consensus mechanisms* serve as a way to ensure 'shared truth' about the content shared within the distributed network [39]. *Cryptography* refers to the techniques of how information are prepared and to what extent an insight into the information can be achieved via concepts such asymmetric *cryptography* [24]. The Public-Key-Infrastructure (PKI) allows the deposition of digital signatures within the distributed network [38]. According to *Wieninger* the combination of *decentralization* and *consensus mechanisms* contributes to the emergence of *transparency* as a trust-relating characteristic [21]. *Transparency* about the transactions inside a distributed network is an important characteristic to ensure trust due to the mitigation of trust or information asymmetry among the network participants, but also provides a source for exposure in terms of opportunistic behaviour of certain malicious participants [22]. *Immutability* refers to the information stored within the blockchain and prevents subsequent modifications of the content stored within the blockchain [38,23]. Immutability can be influenced by the combination of *consensus mechanisms*, *decentralization* and *cryptography* and the previously mentioned *transparency* [23]. Its unalterability in changing these information ensures trust due to the fact that any attempts to change these information will be detected and can be traced back to the origin of malicious activities [23]. *Data-Integrity* ensures that the data sets shared along the blockchain are complete, consistent and valid, whereas the *process- integrity* which is affected by *data-integrity* can be seen as the ability to IT-processes according to quantity and amount [23]. *Process-integrity* can be assured with the help of predefined execution codes, the so-called 'smart contracts' [24,38].The correlations of the seven trust-inherent characteristics are initially derived from the perspective of intercompany networks [23]. As mentioned in chapter 2, intercompany networks and crowds are assumed to be similar and thus allow the transferability of the seven characteristics within the context of crowdsourcing platforms.

4.2 Linkage to the 5P's (functionality mapping)

The previously mentioned characteristics are now projected on the 5 Ps (cf. section 3.3), using the threefold layer-structure proposed by *WIENINGER* [23] which are the technical-, systemical- and application-based layer. As already mentioned in section 3.3, the *platform* consists of the technical components which are *consensus mechanisms*, *decentralization* and *cryptography* [23]. *Peers* and *principals* are the users or rather the stakeholders of the crowdsourcing platform and thus equal to the participants of such an ecosystem and close to the application layer which consist of the elements *data-integrity* and *process-integrity* as essential trust-inherent characteristics. The *problem* as well as the *performance* are object- and result-oriented parts represented by information revealed and shared in the crowdworking system and thus aligned to the characteristics on a systemic-layer such as *transparency* and *immutability*. Figure 2 illustrates the assignments of the 5 Ps based on the seven trust-inherent characteristics. The mapping allows the derivation of provisional suggestions about the interdependencies between the system features which is part of the subsequent section using the maintenance scenario as an example.

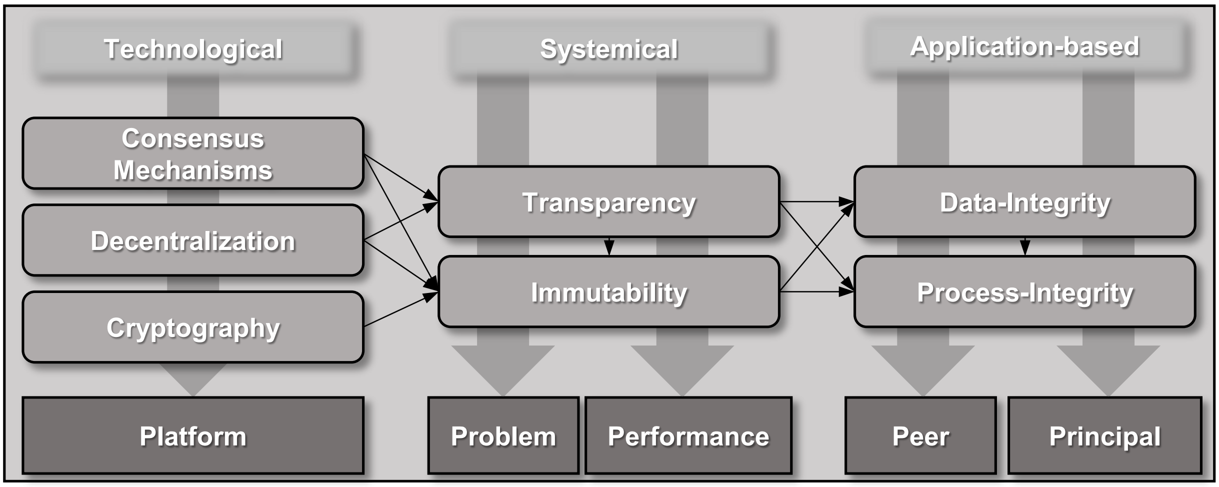


Figure 2: Assignment of the 5Ps onto the trust-inherent characteristics based on [23]

4.3 Use-Case: towards a BCT based crowdsourcing platform for industrial maintenance tasks

To give insights about how the relations illustrated in Figure 2 can be utilised for the industrial service, a short fictional example for demonstration is proposed. Using the example of a maintenance scenario in which an operator of a wind park is receiving unexpected malfunctions of a set of wind turbines which cannot be intercepted by his or her own personal capacity, different sets of competences and tasks are required: spare part installation, disposal of the removed components, electricity handling etc. Due to that the wind park operator (principal) decided to request additional workforce from a crowdsourcing *platform*. Challenges occur due to the lack of transparency about the workers (*peers*) quality and the reliability, so a transparent platform is required to ensure trust into the cooperation. The *problem* is specified by the information about services which are required (type, quantity, time), and the *performance* is determined deadline compliance.

As mentioned in section 4.2, BCT offers a decentralized sharing of the *problem* description to all participants of the crowd. The *process-integrity* can be assured via smart contracts which allows a *decentralized* and automatic fulfilment of transactions based on pre-defined conditions [29]. Following *DING ET AL.*, the use of credit scores and arbitration mechanisms can serve an evaluation of the performance of the tasks fulfilled by the participants (*peer*) and gives a "*set of reasonable evaluation criteria for users*" [29, p. 491]. In addition, *SCHÜTZ ET AL.* propose the use of reputation mechanisms achieved by a mixture of smart contracts on a functional layer and consensus among the participants to ensure trust along the participants in such a network [26]. *Process-integrity* can be assured by smart contracts. The asymmetric *cryptography* ensures that only permitted user can get access to the information shared from the *principal*. This secures the *principal* from being oppressed or exploited by the *peers*. The tamper-proofed blockchain storage ensures valid data storage inside the blockchain history and ensures that malicious attempts are not falsifying the basis for *performance* evaluation. *Process-* and *data-integrity* ensures that the wind park operator can proof whether the qualifications of the applicants are based on valid data which were previously ensured by *consensus mechanisms*. The *peers* in return can ensure that they are not exploited by the *principal* because the BCT as a neutral platform [38] is regulated by the network participants which are represented by the crowd. Malicious participants can be detected and punished [29]. Even the *peer* gain trust in the performance evaluation done by the *principal* because the transactions are validated by the network through *consensus mechanisms* [24,38]. Summarized, BCT offers both the *peers* and the *principal* a sufficient trust basis.

5. Conclusion

To reduce complexity, crowdsourcing platforms offer plenty of possibilities for the organization to work in a decentralized manner. However, challenges such as information asymmetry hamper the success, coordination effort and moreover the trust among participants within such an ecosystem. With the intention of coping with the challenges in terms of information asymmetry along the participants of such a crowdsourcing ecosystem, the BCT must receive more attention. Based on past research, five essential features of a crowdsourcing system were derived and proposed in this paper (5 P s): *Platform*, *Principal*, *Peer*, *Problem* and *Performance* (see Figure 1). In order to show to what extent trust-inherent functionalities can be adapted, seven blockchain-inherent trust-building characteristics were derived from contributions on intercompany networks in association with the BCT are mapped in Figure 2, followed by a projection on the fields of industrial services. Designers of such a crowdsourcing concept can use these assignments as a basis for deriving concrete solutions for each system feature. In this context, also socio-technical design criteria such as: adaptivity, transparency, complementarity, holism, polyvalence and decentralization [40] need to be regarded as core elements in the case of implementing the BCT. Therefore a socio-technical view on platform ecosystems is the basis for a proper implementation of the BCT in such environments [41]. While implementing BCT criteria relating to business conduct (*problem* and *performance*) have to be taken into account as well as criteria relating to stakeholders (*peer*, *principal*, *platform*) and their mutual working relationships (fair pay, working conditions, contracts, management processes and fair participation) [35,42]. Future research should also aim at elaborating concrete blockchain-solutions within the crowdsourcing environment, followed by the derivation of metrics to measure and analyze the cause-effect relationships of implemented functionalities. In the future, the digital world of work will rely on decentralized, heterogeneous and flexible working relationships which will develop in the direction of a so called "crowd economy" [43].

Acknowledgements

The work presented in this paper was funded by the 'Deutsche Forschungsgemeinschaft (DFG) – 276879186/GRK2193'.

References

[1] Botthof, A., 2015. Zukunft der Arbeit im Kontext von Autonomik und Industrie 4.0, in: Botthof, A., Hartmann, E.A. (Eds.), Zukunft der Arbeit in Industrie 4.0. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–8.

[2] Brynjolfsson, E., McAfee, A., 2014. The second machine age: Work, progress, and prosperity in a time of brilliant technologies. Norton, New York, NY, 306 pp.

[3] Henke, M., Heller, T., Stich, V., 2019. Smart Maintenance - Der Weg vom Status quo zur Zielvision (acatech STUDIE). (acatech STUDIE).

[4] Evangelista, R., Guerrieri, P., Meliciani, V., 2014. The Economic Impact of Digital Technologies in Europe. Economics of Innovation and New Technology 23 (8), 802–824.

[5] Baethge-Kinsky, V., Marquardsen, K., Tullius, K., 2018. Perspektiven industrieller Instandhaltungsarbeit. WSI 71 (3), 174–181.

[6] Richter, T., 2017. Industrie 4.0 - Risiken und Chancen für die Berufsbildung -: Betriebliche Weiterbildung als Antwort auf die Implementierung von Industrie 4.0. wbv; W. Bertelsmann Verlag; ProQuest E-Book Central, Bielefeld, Ann Arbor, Michigan, 304 pp.

[7] Schenk, M. (Ed.), 2016. 9. Fachtagung »Anlagenbau der Zukunft«: 19. IFF-Wissenschaftstage 22.-23. Juni 2016.

[8] Mrass, V., Peters, C., 2017. Crowdworking-Plattformen in Deutschland. Working Paper Series 16, Kassel/Germany. https://www.alexandria.unisg.ch/250739/1/JML_615.pdf.

[9] Leimeister, J.M., Zogaj, S., 2013. Neue Arbeitsorganisation durch Crowdsourcing: Eine Literaturstudie. Arbeitspapiere 287, Düsseldorf. https://www.boeckler.de/pdf/p_arbp_287.pdf. Accessed 26 January 2021.

[10] Howe, J., 2006. The rise of crowdsourcing. Wired magazine 14 (6), 1–4.

[11] Leimeister, J.M., 2012. Crowdsourcing, Crowdfunding, Crowdvoting, Crowdcreation. 388-bis 392. St. Gallen, CH: Zeitschrift für Controlling und Management, 388–392.

[12] Siebert, H., 2010. Ökonomische Analyse von Unternehmensnetzwerken, in: Sydow, J. (Ed.), Management von Netzwerkorganisationen. Gabler Verlag, Wiesbaden, pp. 7–27.

[13] Williamson, O.E., 1991. Comparative economic organization: The analysis of discrete structural alternatives. Administrative science quarterly, 269–296.

[14] Aderhold, J., 2005. Unternehmen zwischen Netzwerk und Kooperation, in: , Modernes Netzwerkmanagement. Springer, pp. 113–142.

[15] Clement, R., Schreiber, D., Bossauer, P., Pakusch, C., 2019. Internet-Ökonomie. Springer Berlin Heidelberg, Berlin, Heidelberg.

[16] Cai, R., 2004. Trust and transaction costs in industrial districts.

[17] Spremann, K., 1990. Asymmetrische information. Zeitschrift für Betriebswirtschaft 60 (5/6), 561–586.

[18] Eisenhardt, K.M., 1989. Agency theory: An assessment and review. Academy of management review 14 (1), 57–74.

[19] Schweer, M.K.W., 2012. Vertrauen als Organisationsprinzip in interorganisationalen Kooperationen, in: Schilcher, C., Will-Zocholl, M., Ziegler, M. (Eds.), Vertrauen und Kooperation in der Arbeitswelt. VS Verlag für Sozialwissenschaften, Wiesbaden, pp. 103–121.

[20] Akerlof, G.A., 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. The Quarterly Journal of Economics 84 (3), 488.

[21] Clement, R., Schreiber, D., Bossauer, P., Pakusch, C., 2019. Intermediation: Direkte und indirekte Verknüpfung von Angebot und Nachfrage, in: Clement, R., Schreiber, D., Bossauer, P., Pakusch, C. (Eds.), Internet-Ökonomie. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 153–204.

[22] Notheisen, B., Weinhardt, C., 2019. The blockchain, plums, and lemons: Information asymmetries & transparency in decentralized markets.

[23] Wieninger, S., 2020. Vertrauen in Unternehmensnetzwerken durch Blockchain-Technologie. Apprimus Wissenschaftsverlag.

[24] DIN Deutsches Institut für Normung E.V. DIN SPEC 16597 - 2018: 02: Termonology for blockchains; Text in English. Beuth Verlag, Berlin.

[25] Wang, S., Taha, A.F., Wang, J., 2018. Blockchain-assisted crowdsourced energy systems, in: 2018 IEEE Power & Energy Society General Meeting (PESGM), pp. 1–5.

[26] Schütz, A.E., Fertig, T., Weber, K., Vu, H., Hirth, M., Tran-Gia, T., 2018. Vertrauen ist gut, Blockchain ist besser – Einsatzmöglichkeiten von Blockchain für Vertrauensprobleme im Crowdsourcing. HMD 55 (6), 1155–1166.

[27] Rhyn, M., Blohm, I., 2017. Combining collective and artificial intelligence: towards a design theory for decision support in crowdsourcing.

[28] Abbas, Y., Martinetti, A., Moerman, J.-J., Hamberg, T., van Dongen, L.A., 2020. Do you have confidence in how your rolling stock has been maintained? A blockchain-led knowledge-sharing platform for building trust between stakeholders. International Journal of Information Management 55, 102228.

[29] Ding, Y., Chen, Z., Lin, F., Tang, C., 2019. Blockchain-based credit and arbitration mechanisms in crowdsourcing, in: 2019 3rd International Symposium on Autonomous Systems (ISAS), pp. 490–495.

[30] Vogl, E., 2018. Crowdsourcing-Plattformen als neue Marktplätze für Arbeit: Die Neuorganisation von Arbeit im Informationsraum und ihre Implikationen, 1. Auflage, 11 Abb ed. Rainer Hampp Verlag, Mering, 120 pp.

[31] Leimeister, J.M., Zogaj, S., Durward, D., Blohm, I., September/ 2015. Arbeit und IT: Crowdsourcing und Crowdwork als neue Arbeits- und Beschäftigungsformen. Gute Arbeit und Digitalisierung Prozessanalysen und Gestaltungsperspektiven für eine humane digitale Arbeitswelt, Berlin.

[32] Albert, A., 2017. Kombination von Field Service Management-Apps mit Crowd-Diensten. https://blog.mila.com/de/2017/02/21/combining-field-service-management-apps-with-crowd-service/. Accessed 8 February 2021.

[33] Weinzierl, S., 2020. Per Crowd Service dem Kunden Echtzeit-Hilfe bieten. https://www.instandhaltung.de/instandhaltung-4-0/per-crowd-service-dem-kunden-echtzeit-hilfe-bieten-104.html. Accessed 9 February 2021.

[34] SAP, 2018. Understanding Crowd Service: Leveraging the Gig Economy for Greater Field Service Agility and Better Customer Experiences. https://www.sap.com/products/crowd-service.html. Accessed 8 February 2021.

[35] Benner, C. (Ed.), 2015. Crowdwork - zurück in die Zukunft?: Perspektiven digitaler Arbeit. Bund-Verl., Frankfurt am Main, 420 pp.

[36] Bormann, S., 2017. Crowdwork Erfahrungen von ver.di-Mitgliedern mit der Arbeit auf Online-Plattformen, Berlin. https://innovation-gute-arbeit.verdi.de/++file++5970616824ac060330fe6df7/download/Fact-Sheet_Crowdwork_final.pdf. Accessed 10 February 2021.

[37] Venschott, M., Vogl, H., 2020. Beschaffung technischer Dienstleistungen am Beispiel Instandhaltung: Auch vermeintlich einfache Tätigkeiten können zum Desaster führen. Business & Management, IUBH Discussion Papers No. 4/2020. IUBH Internationale Hochschule, Erfurt.

[38] Fridgen, G., Guggenberger, N., Hoeren, T., Prinz, W., Urbach, N., Baur, J., Brockmeyer, H., Gräther, W., Rabovskaja, E., Schlatt, V., others, 2019. Chancen und Herausforderungen von DLT (Blockchain) in Mobilität und Logistik.

[39] Zhang, P., Schmidt, D.C., White, J., Dubey, A., 2019. Consensus mechanisms and information security technologies. Advances in Computers 115, 181–209.

[40] Hirsch-Kreinsen, H., 2020. Digitale Transformation der Arbeit: Entwicklungstrends und Gestaltungsansätze, 1. Auflage ed., 164eiten.

[41] Kapoor, K., Ziaee Bigdeli, A., Dwivedi, Y.K., Schroeder, A., Beltagui, A., Baines, T., 2021. A socio-technical view of platform ecosystems: Systematic review and research agenda. Journal of Business Research 128, 94–108.

[42] Fairwork, 2020. Fairwork Germany Ratings 2020: Arbeitsstandards in der Plattformökonomie.Großbritannien. Fairwork, Berlin, Deutschland, Oxford, Großbritannien. https://fair.work/wp-content/uploads/sites/97/2020/05/German. Accessed 18 February 2021.

[43] Deutscher Crowdsourcing Verband e. V. Kompetenz für crowdbasierte Geschäftsmodelle. https://www.crowdsourcingverband.de. https://www.crowdsourcingverband.de. Accessed 23 March 2021.

Bibliography

M.Sc. Nick Große (*1992) has been research assistant of the Chair of Enterprise Logistics (LFO) at the University Dortmund since 2019 (Prof. Dr. Michael Henke). He is a member of the research training group ‘Adaption Intelligence of Factories in a Dynamic and Complex Environment’ (GRK 2193).

M.A. Martin Eisenmann (*1987) is a junior researcher at the Social Research Center (SFS) at the University Dortmund since 2019 (Prof. Dr. Jürgen Howaldt). He is also a member of the research training group ‘Adaption Intelligence of Factories in a Dynamic and Complex Environment’ (GRK 2193).

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Use Of A Robot For Parts Provision In Manual Small Parts Assembly

M. Krampe[1], L. Penczek[1], B. Kuhlenkötter[1]
[1]Lehrstuhl für Produktionssysteme, Ruhr-Universität Bochum, Germany

Abstract

Despite the growing popularity of automation, many companies still use manual assembly within their production. Especially when small deviations in product series are involved, automation is usually not flexible enough and does not add value.

This paper describes a new robotic assistance system with the aim to support the worker by separating him from logistics and the choice of the next component. The robot should be the interface between manual assembly and logistics, it will grip the required component and deliver it to the worker. This results in a constellation where the assembly worker can focus on value-added assembly activities and work in a more ergonomic workspace, as there are no complex and overwhelming Kanban shelfs. The consistent use of storage containers means that the robot has no workplace where it must be flexible due to varying products. Furthermore, the logistics area does not have to be adapted to the assembly worker. This enables more efficient logistics solutions that would not be used in a normal manual assembly, such as shelves filled to the ground. In this context, the main challenge is to develop an efficient relationship between the worker and the robot. Here, communication possibilities between the worker and the robot must be investigated, in particular regarding the process flow and error management. Before we focus on new logistic software tools to automate the whole logistic process and to generate the code for the robot, we first want to show the efficiency of our concept by implementing it in form of a demonstrator and testing it on a selected industrial product. Afterwards we want to compare the efficiency and ergonomics of this new approach with the usual manual assembly process.

Keywords

Assistant robotics; assembly; logistics; ergonomics

1. Introduction

Despite the distribution of automation solutions, manual assembly is still firmly anchored in industrial production. This is particularly due to the trend towards increasingly individual products, which require flexible production with small batch sizes. The skill and low training time of human assemblers is still superior to the design effort and programming of automation solutions for many frequently changing products. To deal with the high degree of flexibility required in the design of assembly processes, modular assembly workstations have already become established. These can be virtually assembled by most manufacturers and adapted to the different processes [1]. In industry, a compromise has formed between manual assembly and automation called hybrid assembly. In hybrid assembly, manual assembly is supported by the integration of smaller automations. This combines the advantages of both variants [2]. Despite automation possibilities, this leads to strong research interests in manual assembly, so that the added value can be further increased. Topics such as ergonomics and lean management often come into sharper focus, as they can reveal new potential for improvement. In addition, there are also many efforts to use new robotic technologies such as cobots to support the assembler [3].

2. State of the art

There are many industry efforts in the area of hybrid assembly. In addition to the general improvement of assembly stations through a more variable and ergonomic design, there are various technologies that are intended to increase value added [4]. A pick-by-light or pick-by-projection, for example, is intended to shorten the search time by providing visual assistance [5]. Furthermore, there are body-bound solutions such as smart glasses for augmented reality applications [6]. In the field of assistance robotics, the aim is for humans to be able to work more closely together with the robot. This results in high safety precautions with the associated costs, as well as limited speeds and payloads for the robot. Representatives of these robot systems are, for example, the YuMi from ABB [7] and the iiwa from KUKA [8]. There are also already approaches to use robots in the parts supply of the assembly. A research project of the company item deals with the integration of a collaborating robot into a manual assembly station [9]. Here, the robot works closely with the worker hands him parts and tools. This close collaboration requires high safety precautions whereby the robot may only move slowly and with little payloads. Another example are the assembly workstations from Rose+Krieger, which allow the implementation of a collaborative robot [10]. In this collaboration the robot is also severely restricted by the necessary safety precautions and cannot be utilized to its capacity. Since collaborative robotics near humans can only use its potential to a limited extent due to speed and force limitations, we would like to take a different approach in this paper to integrate robots in manual assembly.

Looking at warehouse logistics, there are already smart systems that simplify warehouse picking and reduce the associated search times. However, the approach of this paper considers the individual assembly station, which has a pre-picked logistics system (e.g. kanban shelf). Of course, this kanban shelf can also be picked in advance with a system such as the Advanced Pick Station [11].

In the following, we will first introduce our new concept and the theoretical possibilities of this approach. Subsequently, the first demonstrator is explained, with which the basics of the concept were first tested.

3. The concept

In contrast to many other new approaches in the research field of assembly, we still want to separate the robot from the worker as far as possible. This should allow each of the two partners to make the best use of their particular skills. For the worker, this means that he can focus more on the value-adding activities of assembly, while the robot performs repetitive and non-ergonomic tasks. Furthermore, this means that expensive and application-specific safety technology can be eliminated, making the systems less expensive and more flexible in their application.

One non-value-added activity is finding the components for the next assembly step. Particularly with many parts containers, which are made available to the worker on the shelf behind the table, search times can add up during assembly. Especially large shelves due to the number of components, as well as high weights, also reduce the ergonomics of assembly.

For this reason, the robot in this concept should only focus on part provision and always provide the worker with the required components. Our concept can be seen in Figure 1 and consists of the worker's assembly area, the robot area, and the logistics system. Furthermore, there is an area for part transfer, where the robot provides the component containers. The robot should only serve as an interface between assembly and logistics and be integrated as little as possible into the assembly process.

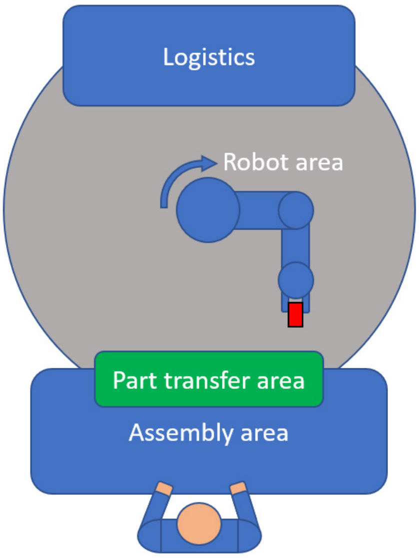


Figure 1: The basic concept

To simplify the gripping process and to be able to use the concept for as many applications as possible, no individual components but always component containers are gripped. This makes it easier to use the system for different assembly processes. The gripped containers are deposited by the robot in the part transfer area. There is enough space for several containers at the same time, so that work steps requiring several components can also be carried out without waiting times. Furthermore, to avoid waiting times, components can already be placed ready for the following work steps. The aim is to reduce the worker's search times by requiring him to select from only five or six ergonomically optimally positioned containers instead of 20 to 60 as we have seen it before in some companies.

To implement this, the robot must know which container is required for which work step. It is also possible for a container to remain with the operator for several steps in succession or for a step to be performed later. This information must already be available before assembly and, in the best case, is already created during product development or production planning.

Communication between the operator and the robot can be implemented most easily by means of a confirmation button, which is pressed after a successful assembly step. For error management, it must also be possible to quickly retrieve boxes that have already been used if the worker has removed too many or too few components.

This concept also opens new possibilities for the logistics system. Since it no further needs to be adapted to the worker, an ergonomic shelving system is no longer required. For example, the effective radius of the robot can be further exploited by filling the rack more to the top or to the bottom. The whole concept can be scaled up as required by using additional linear axes for the robot. If the individual assembly steps take more time, it is also possible for one robot to supply several workers with component containers, as can be seen in Figure 2. The increase in this workspace illustrates the benefit of using faster non-collaborative robots for this task.

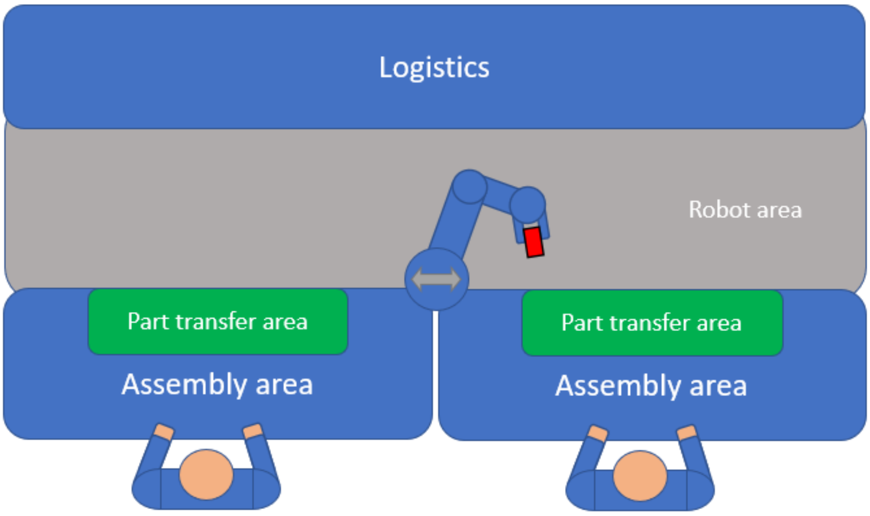


Figure 2: Concept with several stations

In summary, the concept offers the opportunity to reduce the worker's search times while improving the ergonomics of assembly. In the following, a first demonstrator was implemented based on this concept to investigate some initial fundamental questions. How does it affect the assembly process? Can the robot transport the parts to and from the workstation so quickly that the worker does not have to wait? What problems can arise in the area of part transfer?

4. The first demonstrator

In order to be able to test the concept and answer the first questions, a demonstrator was created which contains the basic functions of the concept. In the following, the construction of the demonstrator is explained first. Then the test execution and the results are presented and finally a conclusion is drawn.

The robot-assisted assembly station consists of an usual table as assembly surface, a robot standing behind it as well as an ordinary kanban shelf (Figure 3). A KUKA iiwa was chosen as the robot. Other robots offer higher payloads and speeds, but it turned out that this makes no difference in our first exemplary process. Due to the weights of our components and the times of the individual assembly steps, the KUKA iiwa was used to almost 100% capacity. Our components had a low weight, the component bins were almost empty and the work steps partly time-consuming. This shows the potential benefit of a non-collaborative robot for this application, as it would not yet have reached its limits in this use case. An additive manufactured gripper was used, which was customized to the containers and enabled the robot to remove the boxes from the shelf and store them again. The part transfer area was covered with an anti-slip mat to prevent the crates from being displaced when the parts were removed. This enabled the robot to find the containers even without a vision system.

Figure 3: The hybrid assembly station

In the manual workstation without the robot, the shelf was placed at the rear edge of the table, as it is widely used in the field of assembly (Figure 4). In both cases, the worker was provided with assembly instructions and the necessary tools on the table.

Figure 4: The manual assembly station

Our example process involves the assembly of a small control cabinet. This consists of 22 different components and requires various assembly steps. This includes screwing with and without tools, inserting cables and installing switching terminals. Some steps are fast while others take much more time or require several components at once.

5. Test execution

The study began by recording the assembly times for manual ($t_{m,raw}$) and hybrid ($t_{h,raw}$) assembly. Measurements were made with four test persons, two of whom performed the assembly in one variant first. Table 1 shows the results of the measurements, which variant the subjects performed first and the difference of both recorded times

Table 1: Assembly times

Test person	Started with:	Manual: $t_{m,raw}$	Hybrid: $t_{h,raw}$	Difference
1	manual	15:15	12:00	3:15
2	manual	20:00	15:05	4:55
3	hybrid	19:50	20:00	0:10
4	hybrid	14:10	13:25	0:45

In addition, the search times were recorded. It was particularly difficult to measure the very short search times with hybrid assembly, since the test person only had to choose from a maximum of five boxes in the direct field of view. The following Table Y shows the individual search times. The table already shows that the new concept can reduce search times in the assembly process.

Table 2: Search times

Test person	Started with:	Manual: $t_{m,search}$	Hybrid: $t_{h,search}$
1	manual	1:59	0:31
2	manual	1:20	0:20
3	hybrid	2:39	0:39
4	hybrid	1:30	0:28

When evaluating the recorded times, it is noticeable that the hybrid process reduces the search times, and the process times are also shorter except for person 3. Despite a significantly reduced search time, person 3 is faster in the manual process. This may be due to a learning effect of the test persons, whereby they are faster the second time they perform the assembly, regardless of the variant. This learning effect can be calculated by the following formula.

$$t_{learning} = |(t_{m,raw} - t_{m,search}) - (t_{h,raw} - t_{h,search})| \qquad (1)$$

Using this formula, the following learning effects, shown in Table 3, can now be determined for the four test persons.

Table 3: Learning effects

Test person	Started with:	Learning effect: $t_{learning}$
1	manual	1:47
2	manual	3:55
3	hybrid	2:10
4	hybrid	0:17

If these times are now subtracted from the times of the first performed variant of each test person, these new times can be compared without the influence of a learning effect. In addition to the learning effect, other external influences can of course also play a role which could lead to the difference in times. However, according to the test subjects, this learning effect should have the greatest influence. The now processed times are shown in Table 4.

Table 4: Assembly times without the learning effect

Test person	Started with:	Manual: $t_{m,l}$	Hybrid: $t_{h,l}$	Difference
1	manual	13:28	12:00	01:28
2	manual	16:05	15:05	01:00
3	hybrid	19:50	17:50	02:00
4	hybrid	14:10	13:08	01:02

As expected, the differences between the two methods also reflect the differences in search times described above. These results will now be interpreted in the following.

6. Conclusion

Since the search times of the worker could be minimized with the help of the robot, the total assembly time was also reduced by just this amount. In relation to our application example, this proportion is only very small. The special opportunity of this approach is in its scalability. Even if the assembly process becomes more complex and requires more part containers, nothing changes for the worker. He still only has to choose between the few containers provided by the robot. This can pay off especially in a production with many variants, when the search times due to the changing containers cannot be reduced by a learning routine in the assembly process.

Furthermore, this makes it easier to have all the parts containers for a variety of different assembly processes at one station and thus to be able to switch more flexibly between the variants. It makes no difference to the worker's workload whether he assembles something consisting of 20 or 200 different components. The ergonomics of the workstation, consisting of the gripping ranges and the search times, are always the same, and the worker can focus more on the assembly process itself.

The previous investigation was carried out with a KUKA iiwa, and the robot's speeds and payloads had no negative effects in the selected application. The worker did not have to wait for the robot at any time, despite an assembly with very different work steps. As a result, it can be expected that a more powerful robot will also be able to supply the assembly with components quickly enough with even shorter process steps. Furthermore, an ordinary industrial robot is also significantly less expensive, which means that the changeover to the new workplace pays off financially at an earlier stage. The possibility of supplying several workers with components is also still an option if the process permits this. The next problem arises from the part transfer between the robot and the worker. The problem of locating the boxes for the return to the rack was solved in the demonstrator by simple anti-slip mats. For the industrial application, another possibility must be found here. A vision system would be possible, which could also be used for safety. When a powerful industrial robot is used, a solution must be found for the safety of the worker. The interface between man and robot is the transfer of parts through the containers. The area is therefore very small and direct contact could be avoided. A vision system could be used to track the worker so that the robot knows at what speed it is allowed to approach the area or if it should even stop briefly. A container with high weight or sharp objects, which is moved at high speed, represents the biggest safety risk, for which a solution must be found in following studies.

References

[1] Lotter, B., Deuse, J., Lotter, E., 2016. Die Primäre Produktion: Ein praktischer Leitfaden zur verlustfreien Wertschöpfung. Springer Vieweg, Berlin.

[2] Lotter, B., Wiendahl, H.-P., 2012. Montage in der industriellen Produktion: Ein Handbuch für die Praxis, 2nd ed. Springer Vieweg, Berlin.

[3] Beumelburg, K., 2005. Fähigkeitsorientierte Montageablaufplanung in der direkten Mensch-Roboter-Kooperation, Stuttgart.

[4] Winkler, H., Zinsmeister, L., 2019. Trends in digitalization of intralogistics and the critical success factors of its implementation. BJO&PM 16 (3), 537–549.

[5] Funk, M., Lischke, L., Mayer, S. Neue Impulse für visuelle Kommissionierassistenssysteme aus der Mensch-Computer Interaktion: Ein Überblick über die Kommissionierassistenz im Zeitalter von Industrie 4.0. Warehousing 4.0 2017.

[6] Hüer, L., Zobel, B., Birkel, H., Thomas, O., 2020. State-of-the-Art von Smart Glasses zur Unterstützung von Logistikprozessen, in: Thomas, O., Ickerott, I. (Eds.), Smart Glasses. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 51–68.

[7] ABB Automation GmbH. YuMi - IRB 14000. https://new.abb.com/products/robotics/de/kollaborative-roboter/yumi. Accessed 26 May 2021.

[8] KUKA AG. LBR iiwa. https://www.kuka.com/de-de/produkte-leistungen/robotersysteme/industrieroboter/lbr-iiwa. Accessed 26 May 2021.

[9] item Industrietechnik GmbH, 2019. Mensch-Maschine-Kollaboration für die manuelle Montage. https://blog.item24.com/digitalisierung/mensch-maschine-kollaboration-fuer-die-manuelle-montage/. Accessed 25 May 2021.

[10] RK Rose+Krieger GmbH. Montagearbeitsplätze mit COBOT-Anbindung – kollaborierende Roboter unterstützen den Prozess. https://www.rk-rose-krieger.com/deutsch/produkte/modul-technik/arbeitsplatzsysteme/arbeitsplaetze-mit-cobot-anbindung/. Accessed 25 May 2021.

[11] SSI SCHÄFER FRITZ SCHÄFER GMBH. Pick-to-Tote-Arbeitsplätze. https://www.ssi-schaefer.com/de-de/produkte/kommissionieren/halbautomatisches-kommissionieren/advanced-pick-station-9946. Accessed 31 May 2021.

Biography

Michael Krampe (*1993), M.Sc., works as a research assistant at the Chair of Production Systems (LPS) at the Ruhr-University Bochum and conducts research in the field of industrial robotics.

Lars Niklas Penczek (*1993), M.Sc., works as a research assistant at the Chair of Production Systems (LPS) at the Ruhr-University Bochum and conducts research in the field of industrial robotics.

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.

2nd Conference on Production Systems and Logistics

The Impact Of Manufacturing Execution Systems On The Digital Transformation Of Production Systems - A Maturity Based Approach

Markus Fischer[1], Günther Schuh[1], Volker Stich[1]

[1]*Institute for Industrial Management (FIR), Aachen, Germany*

Abstract

With the focus of manufacturing companies on the digital transformation, Manufacturing Execution Systems are market-ready, modular software solutions for manufacturing companies to integrate the value-adding and supporting processes horizontal and vertical in the company. Companies, especially small and medium-sized companies, face high internal and external costs for the implementation of the MES modules. An advantage of MES is the possibility to implement the systems in a continually, module-by-module approach, with the benefit of timely distributed investments. By realizing fast improvements, companies can use the benefits for further module implementations. This paper proposes a maturity model to measure the impact of an MES on the digital transformation of the company's production systems. The model fulfils two purposes. The first, companies can measure the impact based on the difference between its current maturity index and the potential index of an implemented MES. The second is, the user can identify what impact an MES has in general on the digital transformation since the developed maturity model is derived from an established industry 4.0 maturity model. The development of the maturity model is based on the methodologies of AKKASOGLU and focuses on the further development of an established model. As an outlook, the application of the model will be described briefly. The proposed maturity model can directly be used by practitioners and offers implications for further development of MES functionalities.

Keywords

Digitalization; Manufacturing Execution System; MES; Maturity Model, Maturity Index;

1. Introduction

Digitalization and horizontal and vertical integration promise to be a way of responding to the challenges of reducing costs, increasing quality, shortening delivery times and fluctuating and difficult-to-predict customer demands. The goal of digitalization is to integrate the physical and digital worlds to accelerate decision-making and business operations. [1] The goal is to empower companies to learn from data. In production, this means in concrete terms that near-real-time data and information are used to create a control loop with production planning and control. [2]

Companies, which are dealing with digital transformation in production, must concern the implementation and use of Manufacturing Execution Systems (MES). Currently available MES cover many of the requirements of a digitalized production system. More than capturing and distributing data in real-time, MES are capable of connecting different areas of the production and the supporting processes to enable the company to build a digital shadow of their production system and the processes within. [3] Digital shadows are digital models of the relevant business processes of the company filled with in real time acquired data which are used for analyses and data based decision making [4] Therefore, MES are a relevant part of realizing the digital transformation in production companies.

DOI: https://doi.org/10.15488/11258 publish-Ing.

During the implementation of an MES, companies face many challenges. One of the biggest hurdles, especially for small and medium-sized companies, are the high internal and external costs for the implementation. [5] To tackle this hurdle companies can spread costs over a longer period with a successive implementation. This time can cover from two up to more than five years based on the number of implementation iterations and factory locations in which an MES has to be implemented. Due to the modular structure of MES, companies have to choose which functions and modules to implement first concerning the highest benefits within the successive implementation approach. Here companies face the unsolved challenge to estimate the benefits of the MES implementation ex-ante. MES are complex systems due to their extensive functions and modular structure. The effects of implementation on the various areas affected are difficult or impossible to assess due to the systems complexity.

To address this challenge and support companies, this paper presents a way to assess the impact of an MES implementation with a maturity model. Maturity models represent a relevant assessment tool in the literature and practice to make complex issues easily assessable. In recent time, many maturity models were presented which help companies to assess their status-quo in digitalization and plan their next steps. Since there is no maturity model, for assessing the impact of an MES implementation, a new maturity model has to be derived. [6–8]

2. State of the Art on the assessment of the impact of implementations

This section presents the state of the research in the fields of impact and benefit assessment for implementations of IT systems especially for business applications in production management.

The ex-ante assessment of the impact is a well-known field in the research of information systems and business software. Nevertheless, no approach truly fulfils all needs of practitioners. The difficulties are measuring tangible and intangible benefits and combining both categories in an easy-to-use method. In this section, the presented approaches can be divided into system-specific and general approaches. The general approaches present models to assess IT performance and the investments in IT systems. These approaches are not specific for a kind of IT systems (e.g. MES). The general models need to be adapted by the companies before the usage and they need IT system specific knowledge for the functionalities and the expected impact of the system. Therefore, these models mostly lack the practicability for companies to easy measure the impact of an MES due to the effort of transferring the models for the specific use case. [9,10] In [9] a model to assess the direct effects of IT on the efficiency of the production planning and control is presented, while in [10] the HÄNSCH presents an approach to monetarize the effects of intangible benefits.

MES-specific approaches tend to address the mentioned challenge. KLETTI presents a tool to calculate the Return of Invest of a MES implementation. This calculator focuses on some specific areas where it is easy to calculate the benefits in terms of reduces cost, stocks or rework. [11] With this approach, KLETTI neglects different areas and intangible benefits of the MES. OBERMAIER AND KIRSCH present a paper, which addresses the benefits of an MES implementation in the area production by a comparison of three different KPIs before and after the implementation. [12] Their paper delivers a first quantified assessment/evaluation of the impact of an implementation, but it is not easily transferable to other companies due to different status-quos before the implementation. The VEREIN DEUTSCHER INGENIEURE presents a norm regarding the costs and benefits of an MES. The norm gives a holistic overview the tangible and intangible benefits. [5] It lacks a methodology to assess the benefit of implementation quantified but it describes the effect chains between MES modules and the benefits.

3. Research methodology

To derive a maturity model to measure the impact of an MES, the methodology by AKKASOGLU will be used. Akkasoglu proposes a nine stepped approach to derive a maturity model that is shown in Figure 1. The first three steps are for preparation. The first three steps cover the determination of the need and the specification of the requirement as well as the analysis of existing maturity models. [13] These steps were covered by the first two sections of the paper. The focus of the paper lies on the construction phase that consists of four steps to set-up the maturity model. These steps will be covered in this paper. The first step of the construction phase covers the design of a reference model to describe the area of assessment. Therefore, the production system based on the reference model of MEISSNER will be described [14]. The second step covers the deduction of maturity objects. The maturity objects should be significant for assessing the maturity levels of the production system. In the paper, the maturity objects will be derived by selecting the tasks of the reference model that are supported by an MES. The third step of the construction phase is the weighting of the maturity objects. For this paper, there is no need in weighting the objects since every tasks is weighted the same for the construction of the maturity model. The applying company can do a weighing of the tasks individually. The last step is setting-up the maturity model with maturity levels in a matrix. In this paper, the matrix also shows to which maturity level the tasks will be supported by an MES. The last two steps of AKKASOGLU's methodology are summed up in the application phase.

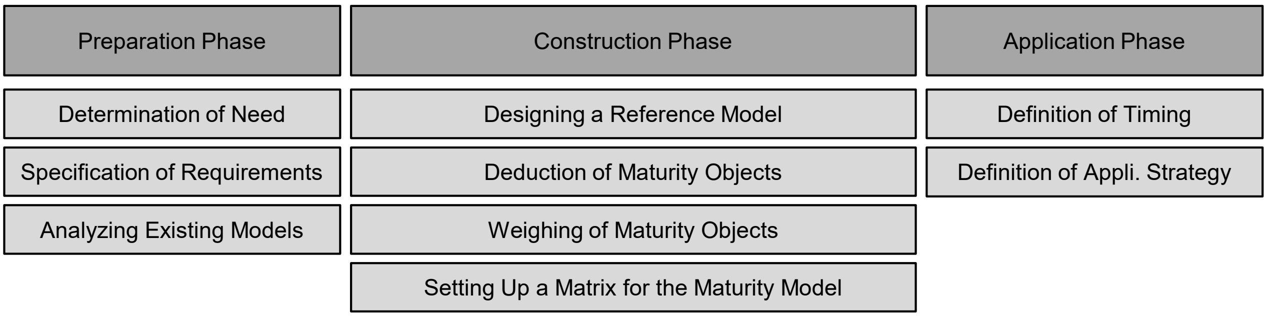


Figure 1: Methodology for designing a maturity model by AKKASOGLU [13]

4. Maturity Model for assessing the impact of a Manufacturing Execution System

In this section, the paper presents the approach and the results of the described steps. The steps for designing the maturity model and assessing the impact of an MES on the production system will be described in the following subsections.

4.1 Design of a reference model

This section deals with the design of the reference model. The reference model should define and mark out the area of investigation. The reference model is a generalized version of the area of investigation and should be detailed enough to be applicable but not so detailed that the effort for the assessment increases unnecessary. [13] In this case, the rough structure for the designed reference model is based on the model for production companies by MEISSNER. MEISSNER structures the production system in seven areas. The areas cover the design and development, industrial engineering, production planning and control, production and assembly, quality assurance, maintenance management as well as production logistics. [14] For each of the areas, standard literature and norms were used to identify the key processes and tasks. The 105 identified key processes and tasks were grouped in 25 task groups. In Figure 2 the process for the design of the reference model is shown. The area of design and development was described based on the VDI 2221 and the 19 tasks are divided into the groups of main, supporting and cross-cutting activities. [15] The area of industrial engineering consists of 11 tasks in of the three groups preparing tasks, the creation of the routing and NC programming. [16] Production planning and control (PPC) is structured based on the Aachen PPC model and structures the total of 30 tasks into eight groups. The groups cover the production program planning, the production demand planning and in-house production planning and control, the planning and control of procurement, inventory management, order management, data management and controlling. [17] The area of production and assembly consist of the group of execution, where the two tasks are divided into value-adding and conditional-value-adding activities. The second group of the area production and assembly contains the two supporting activities of acquisition and provision of information. [18] The area of production logistics consists of five tasks, which are split into the groups of planning and execution. [19,20] The area of quality assurance covers 19 tasks in five groups. The four groups of planning, execution and analysis of quality assurance tasks as well as the test equipment management are located in the quality forward chain. The tasks group complaint management covers five activities and is covering the quality backward chain. [21] The Maintenance management covers 17 relevant tasks in three groups. The groups are divided into the planning part of maintenance management, the execution of the maintenance orders as well as the analysis of acquired data. [22] With the 105 tasks the production system and its tasks and processes are described in a holistic way.

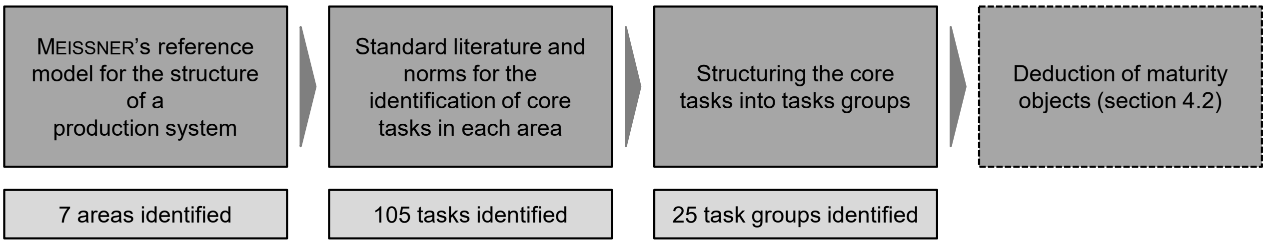


Figure 2: Process for the design of the reference model

4.2 Deduction of maturity objects

The deduction of maturity objects aims for the identification of relevant features of the area of investigation to describe the production system and the impact of the MES on it. Therefore, the papers identifies the tasks, which are supported and influenced by the functionalities of an MES. The structure of functionalities of a general MES is described in the VDI 5600 norm. There, the MES is structured into ten modules with various sub-functionalities. [23] KLETTI describes many functionalities of the modules based on the system of HYDRA, a German MES developed by the company MPDV. [3] With the explanations of the norm and KLETTI, relevant tasks of the above mentioned had to be identified to fill the matrix with the production system's tasks supported by an MES. The method for identification is based on three cases. The first case for support of a task or process is given if the function and the task are named the same. As an example, this case is given for the MES function of data acquisition and the task of data quality acquisition. The second case if the content-related consistency between the MES function and the task. For example, the MES functions of detailed planning are supporting the tasks of setting up an order sequence. The third case represents a partial support of the MES for tasks. Here, information that is used in the corresponding function can be useful for the supported core task. There is neither name- nor content-related consistency on this case. In Figure 3, the identified tasks of production planning and control as well as the support of the MES module form the VDI 5600 are shown. Here it can be seen, that production program planning as well as procurement planning and control are not supported by an MES. The rest of the tasks groups is supported by at least one of the MES modules.

Production Planning and Control (Supports ↓)	MES Modules (VDI 5600)									
	Order Management	Equipment and Maintenance Management	Data Acquisition	Energy Management	Detailed Planning and Control	Information Management	Performance Analysis	Material Management	Personnel Management	Quality Management
Production Program Planning										
Production Demand Planning	●				●					
In-house production planning and control	●	●		●	●	●		●	●	●
Procurement Planning and Control										
Order Management	●					●	●	●		
Inventory Management			●				●	●		●
Controlling	◑	◑	◑	◑	◑	●	●	◑	◑	◑
Data Management	●	●	●	●	●	●		●	●	●

● Case 1 & 2: Tasks are supported by MES
◑ Case 3: Tasks are supported indirectly by MES

Figure 3: Task-Function-Matrix for the tasks of production planning and control

The support of an MES function was checked for each task and summarized by the task groups. With the methodology, 66 of the 105 tasks and 22 of 26 tasks groups were identified as relevant maturity objects to measure the impact of an MES on the production system. The Figure 4 and Figure 5 (Section 4.3) present the identified task groups. Task groups without a support from an MES were left blank.

4.3 Design of the maturity matrix and impact of Manufacturing Execution Systems on the production system

The last step of the construction phase is to set up the maturity model, which will be done in a matrix. The tasks groups are represented as a line in the matrix. The maturity levels will be represented in a column. SCHUH ET AL. presented the acatech Industrie 4.0 Maturity Index as the result of an applied research process in 2017 and it was updated in 2020 with the experience gained while applying it. The acatech Industrie 4.0 maturity Index consists of four key areas with six maturity levels. The four key areas are resources, information systems, organization, and culture. [4] In the further, this maturity model will be used due to its usage and further development and due to the dissertation of SCHMITZ.

The acatech Industrie 4.0 maturity model has six maturity levels. The first two level *Computerization* and *Connectivity* represent the basis of Automation. The following four levels of *Visibility*, *Transparency*, *Predictability* and *Adaptability* represent the levels of Industry 4.0 and the digitalization. The maturity level of *Visibility* describes the capability of a company to have real-time data acquisition of all relevant processes. With the so-called digital shadow companies can analyse what is happening in the processes. The next level describes the capability of a company to analyse the data from the digital shadow to get insights on the correlations between events. Therefore, the company needs diagnostic analytics. Based on the digital shadows and the correlations, the fifth maturity level describes the state in which a company can use the acquired data to make prognostics on future events. This enables companies to construct realistic scenarios for the decision-making. The last maturity level describes the capability of autonomous decision-making by the systems through the evaluation, selection and initiation of measures based on predicted scenarios. [4]

To measure the impact of the MES, an information system, a metric is needed. SCHMITZ presents a metric to measure the maturity levels of information management in his dissertation. He develops a morphology to assess the maturity levels on 11 features. The features are split in the three groups. The first group represents features of the information logistics, which sums up the acquisition and distribution of information. The second group covers features of the data quality, data interfaces and transparency and connectivity of the IT system landscape. The last group covers the features of the IT infrastructure like data management, infrastructure management and the user interface. Each feature is described by forms, which are assigned to a maturity level by means of typologization. The maturity level can be determined by evaluating each feature and its form found in the company. Since information systems carry out the tasks of information management, this paper uses the metric to assess the impact of an MES.

Each task was assessed with the morphology by SCHMITZ and his assessment methodology described in his dissertation. [1] In Figure 4 and Figure 5, the result of the assessment is shown. Dark grey areas show that each assessed task in the group reached the marked maturity level by using an MES, whereas the crossed areas are indicating that an MES supports some tasks to a higher and therefore marked maturity levels. In the following, the figures and the impact are explained.

An MES can assist the support activities within the processes of *design and development* with the diagnostic analysis of the acquired quality data from production processes and the quality controls. With the identified correlations, the design and development can improve the product design in change requests or in new products. The *industrial engineering* get support from the MES by the better integration of data from different systems like ERP, CAD and PLM with data from the production. Especially in the planning of the routing, the acquired production data and its analysis helps the department in defining the processing and set-up times. In addition, the functions from information management help by saving and distributing of NC programs to machines with respect to the production plan.

The tasks and processes of *production planning and control* profit a lot from the usage of an MES. The demand planning and the in-house production planning and control profit from the availability of real-time data for the planning process as well as the analysis functions, which helps to analyse correlation between planning tasks and the latter execution in the production and assembly. The MES functions for fine planning support the production system by providing capabilities for optimization and simulation of the in-house planning and control. These functions enable the simulation of different planning scenarios and the selection of the best production plan. The MES functions of material management enhance the capabilities of the inventory management by giving transparency over the work in progress on the shop floor. The MES realizes tracking and tracing of products along the manufacturing and assembly processes. This also support the tasks of the order management and the controlling of the production system, since the status and progress of the order is tracked almost in real-time. With the data acquisition and analysis functions of the MES, the order management can identify correlations and root causes for deviations and derive reaction measures. The MES supports the data management with improvements in the data quality of the master data and transaction data. Latter data quality is improved by the possibilities of sensor integration and automatic data acquisition as well as functions for plausibility checks while manual inputting data. MES supports the production planning and control in such a way that they reach at least the maturity level of *Visibility*. Some tasks are even enabled to reach the maturity level *Adaptability*.

The tasks of the *production and assembly* are enabled to reach the maturity level of *Transparency*. The MES supports the execution tasks by data acquisition, automatic workflows and the presentation of user and tasks-specific information. It enables the paperless production. Here the MES functions of order management, information management as well as the data acquisition are used.

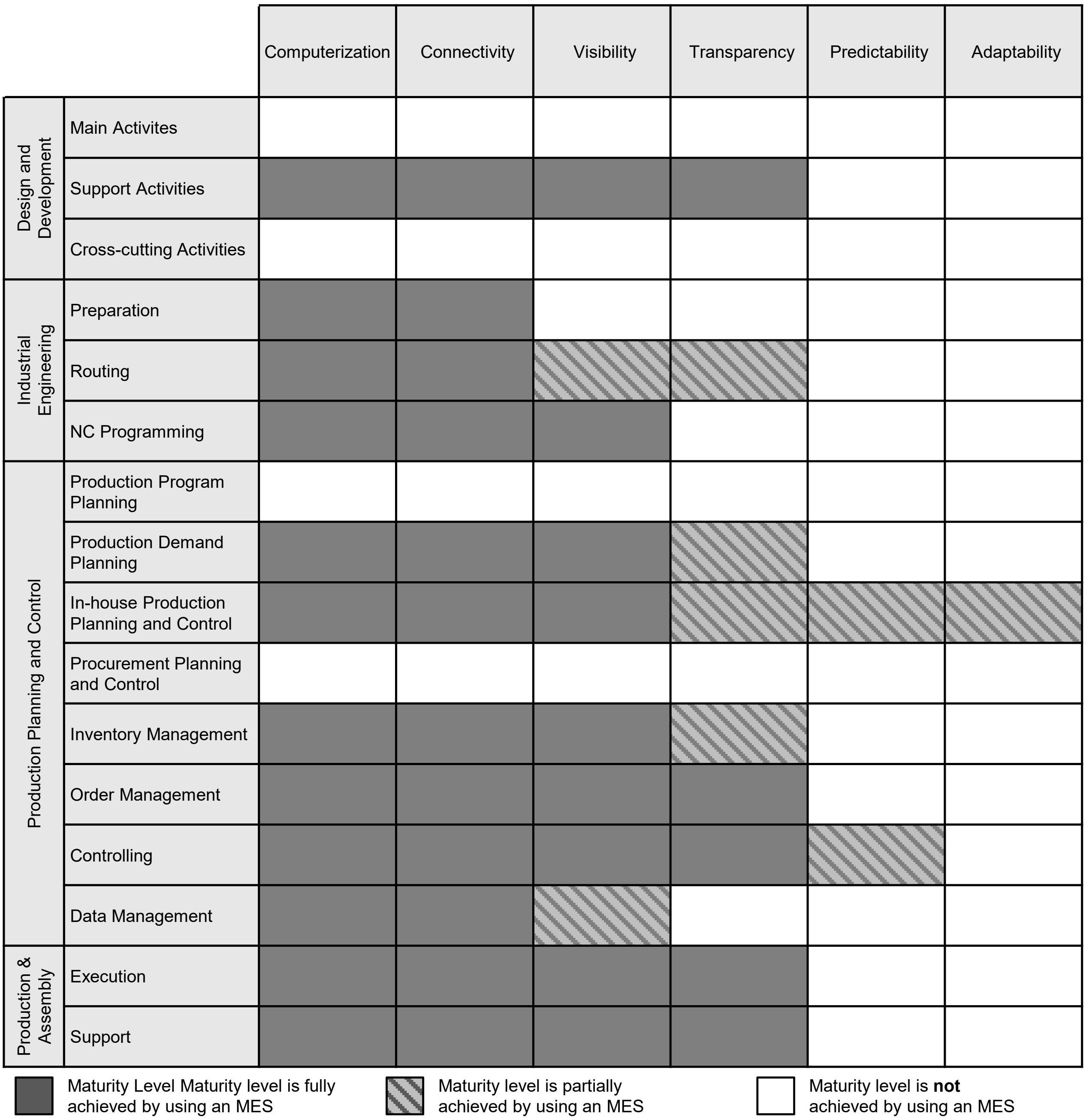


Figure 4: Impact of an MES on the production system (1/2)

In Figure 5, the impact of an MES on the areas of *production logistics*, *quality assurance* and *maintenance management* is shown. The tasks *production logistics* are supported similar to the production processes. The MES has not a big impact on the executional tasks in the logistic area since the MES functions are designed around the manufacturing and assembly processes. Therefore, the focus on the logistics tasks is lesser due the historical development. The planning of the production tasks is connected to the fine planning of production and it provides a synchronisation of manufacturing and transport operations. An MES enables the *production logistic* to reach the *Transparency* level of the maturity model.

The area of *quality assurance* also profits from the MES integration The MES functions of the module quality management enhance all tasks of the area. The quality planning is supported by the MES in the preparation of test plans by the automatic identification of test characteristics from digital component drawings. The execution of the test plans profits by the data acquisition functions and the possibility of an MES to integrate test equipment with interfaces to transfer data. The analysis functions enable not only the analysis of quality data but also its correlation with acquired machine and production data. Through the possibilities of tracking and tracing the tasks of quality data analysis can reach the *Transparency* level. The tasks of test equipment management do not experience any impact in the area of industry 4.0 through the use of an MES. They are supported in such a way that the level of *Connectivity* is reached since the MES takes over repetitive tasks and manages the data of the test equipment without linking it to the digital shadow of the acquired data. The complaint management is supported by an MES to reach a least the level of *Connectivity* since some tasks are just supported by managing data without further support, but the tasks of the error and root-cause analysis can be supported by the analysis functions of an MES. Here, the acquired data is analysed and correlations are drawn. This supports the quality operators in identifying the error and selecting countermeasures.

The area of *maintenance management* profits from the usage of an MES due to the better integration of the maintenance into production. The long-term planning of the maintenance activities profits from the data availability and the insights from identified correlations. An MES enables a strategic shift from time-based to a condition-based maintenance approach. The analysis functions as well as the dedicated maintenance management functions enable the planning tasks to reach maturity levels up to *Transparency*. The controlling and execution profits from the integration to fine planning functions. Here, the planning of planned and unplanned maintenance order is synchronized with the fine planning of production. Since the same functions are used the same maturity level can be achieved. Furthermore, the maintenance operator have access to maintenance instruction and information on MES terminals placed at machines or nearby. Also, they can analyse machine data from machine that need maintenance and can derive implication to the condition. This corresponds to the maturity level of *Transparency*.

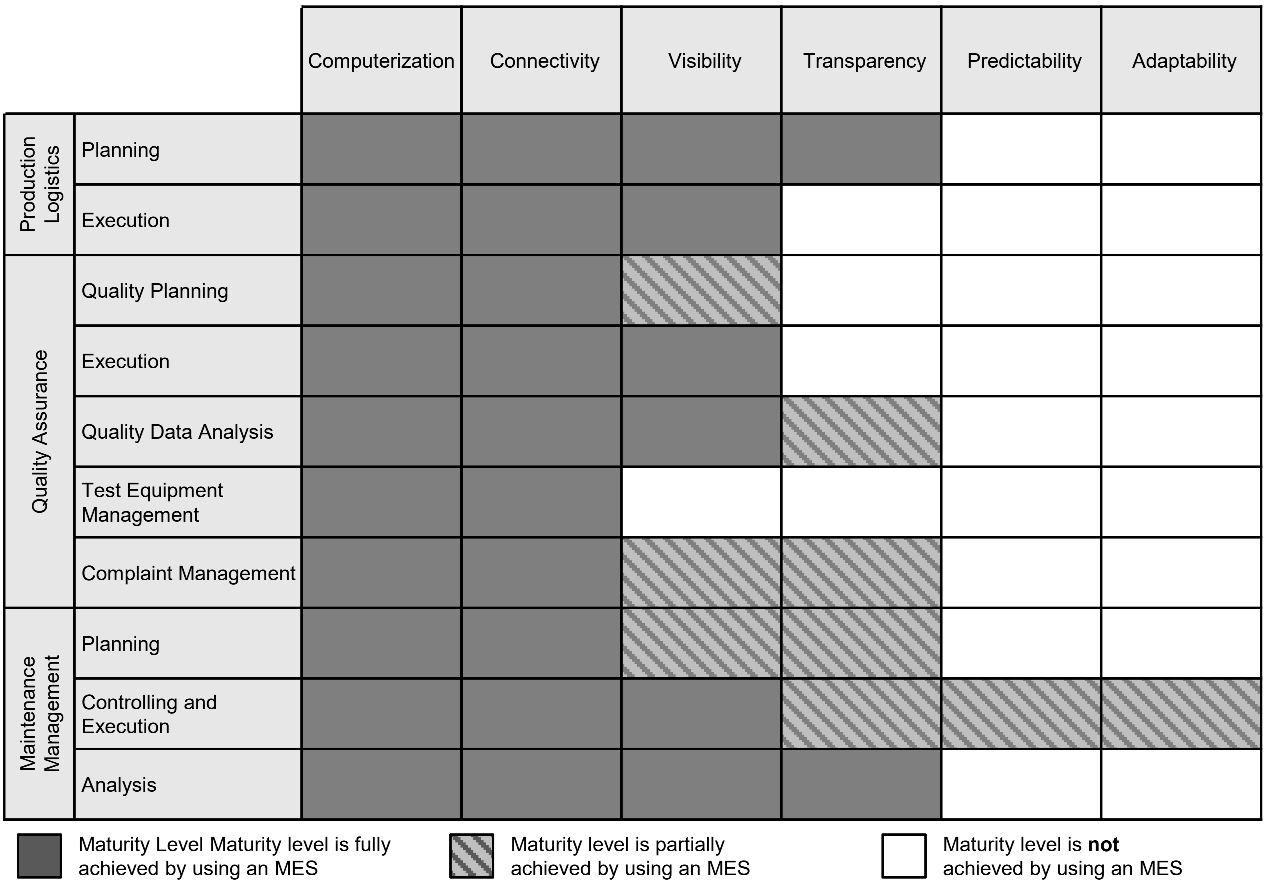


Figure 5: Impact of an MES on the production system (2/2)

5. Conclusion

This paper proposed an adapted maturity model to assess the impact of an MES on the production system. After determining the need for an assessment tool, the acatech Industrie 4.0 maturity model was selected as a basis for the development of a maturity model to measure the impact of an MES. Subsequently, a reference model was built to describe the production system. In the second step, the identified tasks and their groups were reduced accordingly. In the third step, the morphology of SCHMITZ was used in order to have an evaluation logic, which enables the assessment of the impact of an MES on the production system. The impact of an MES on the specific areas of the production system was presented in the maturity matrix. The assessment shows that an MES has different impact on the identified tasks. In many areas, the MES provides support to reach the levels of Visibility and Transparency. Even the higher maturity levels can be reached. With this assessment, practitioners can now assess their own production system area. To do this, they have to determine the maturity level of their company in accordance with the acatech model procedure. By comparing their level with the maturity level supported by the MES, they can identify the areas in the company that would benefit most from an implementation. Relevant MES modules for the high potential areas can be identified with the tasks-function-matrix. In future reserach, based on the model, a procedure can be designed that describes how to apply the models to structure the maturity-based implementation. In addition, the model can be extended for other IT systems that are used in the production. For MES providers, the model can be used to design smart services that support tasks beyond the current state of the art.

Acknowledgements

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

References

[1] Schmitz, S., 2018. Bewertbarkeit des Reifegrads vom Informationsmanagement in produzierenden Unternehmen in Industrie-4.0-Umgebungen, 1. ed. Apprimus Verlag, Aachen, 248 pp.

[2] Wetzchewald, P., 2020. Gestaltungsmodell zur Steigerung der Regelbarkeit von Produktionsplanungs- und -steuerungssystemen: Design model for increasing the regulability of production planning and control systems, 1. ed. Apprimus Verlag, Aachen, 252 pp.

[3] Kletti, J., Deisenroth, R., 2019. MES-Kompendium. 1. ed. Springer, Berlin,, 279 pp.

[4] Schuh, G., Anderl, R., Dumitrescu, R., Krüger, A., ten Hompel, M., 2020. Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten, 64 pp. www.acatech.de/publikationen.

[5] Verein Deutscher Ingenieure, 2013. Fertigungsmanagementsysteme: Wirtschaftlichkeit - Blatt 2. Beuth Verlag, Berlin, 44 pp.

[6] Basl, J., 2018. Companies on the Way to Industry 4.0 and their Readiness. Journal of Systems Integration (3), pp. 3–6. doi:10.20470/jsi.v9i3.351.

[7] Felch, V., Asdecker, B., Sucky, E., 2019. Maturity Models in the Age of Industry 4.0 – Do the Available Models Correspond to the Needs of Business Practice?, in: Bui, T. (Ed.), Proceedings of the 52th Annual Hawaii International Conference on System Sciences (HICSS), Wailea, Hawaii, USA, pp. 5165–5174.

[8] Mittal, S., Khan, M.A., 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, pp. 194–214. doi:10.1016/j.jmsy.2018.10.005.

[9] Dürr, P., 2013. Modell zur Bewertung der Effizienz der IT-Unterstützung im Auftragsabwicklungsprozess von produzierenden KMU. Fraunhofer, Stuttgart.

[10] Hänsch, F., 2015. Finanzwirtschaftliche ex-ante Bewertung intangibler Benefits von IT-Projekten. HMD 52 (6), pp. 945–957. doi:10.1365/s40702-015-0187-4.

[11] Kletti, J., 2007. Konzeption und Einführung von MES-Systemen, 1. ed. Springer, Berlin, 306 pp.

[12] Obermaier, R., Kirsch, V., 2017. Betriebswirtschaftliche Wirkungen digital vernetzter Fertigungssysteme: Eine Analyse des Einsatzes moderner Manufacturing Execution Systeme in der verarbeitenden Industrie, in: Obermaier, R. (Ed.), Industrie 4.0 als unternehmerische Gestaltungsaufgabe, vol. 62. Springer Fachmedien, Wiesbaden, pp. 191–217.

[13] Akkasoglu, G., 2013. Methodik zur Konzeption und Applikation anwendungsspezifischer Reifegradmodelle unter Berücksichtigung der Informationsunsicherheit. Dissertation, Erlangen, 100 pp.

[14] Meissner, J.P., 2018. Adaptives Abweichungsmanagement in der Fertigungssteuerung bei Kleinserien: Adaptive deviation management in production control for small batch production, 1. Auflage ed. Apprimus Verlag, Aachen, 223 pp.

[15] Verein Deutscher Ingenieure, 2019. Entwicklung technischer Produkte und Systeme: Modell der Produktentwicklung. Beuth Verlag GmbH, Berlin 03.100.40, 56 pp.

[16] Eversheim, W., 2002. Organisation in der Produktionstechnik 3: Arbeitsvorbereitung, 4. ed. Springer, Berlin, Heidelberg, 295 pp.

[17] Schuh, G., Stich, V., 2012. Produktionsplanung und -steuerung 1: Grundlagen der PPS, 4. ed. Springer, Berlin.

[18] Eversheim, W., 1989. Fertigung und Montage, 2. ed. VDI-Verl., Düsseldorf, 335 pp.

[19] Lödding, H., 2019. Produktionslogistik, in: Furmans, K., Kilger, C. (Eds.), Betrieb von Logistiksystemen. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 107–131.

[20] Pfohl, H.-C., 2018. Produktionslogistik, in: Pfohl, H.-C. (Ed.), Logistiksysteme. Springer, Berlin, pp. 201–220.

[21] Schmitt, R., Pfeifer, T., 2015. Qualitätsmanagement: Strategien - Methoden - Techniken, 5. ed. Hanser, München, 864 pp.

[22] Pawellek, G., 2016. Integrierte Instandhaltung und Ersatzteillogistik. Springer Berlin Heidelberg, Berlin, Heidelberg, 440 pp.

[23] Verein Deutscher Ingenieure, 2015. Fertigungsmanagementsysteme. Beuth Verlag, Berlin, 40 pp.

Biography

Markus Fischer, M. Sc. (*1992) has been working as a project manager at the Institute for Industrial Management (FIR) at the RWTH Aachen since 2017. In current position as the leader of the research group Production Control as part of the Production Management department, he supports companies in improving processes by the usage of different IT systems. Also, he is working in the research project "Internet of Production", which aims to develop data-driven decision support systems.

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.

2nd Conference on Production Systems and Logistics

Application Of Innovation Diffusions Models In Factory Planning For Fuel Cell Systems

Maximilian Stange[1], Lukas Roth[1], Marian Süße[1], Steffen Ihlenfeldt [1;2]

[1] *Fraunhofer Institute for Machine Tools and Forming Technology IWU, Chemnitz, Germany*
[2] *Institute of Mechatronic Engineering, Technische Universität Dresden, 01069 Dresden, Germany*

Abstract

The planning of factories and production processes is subject to constantly changing requirements and has always been challenged by uncertain future forecasts. Factory planning projects initiated for the production of innovative products are particularly affected. For these products, there is almost no historical or empirical data available, so the forecasts can only be based on estimated influencing variables and data from similar products. Accordingly, the risk that the selected capacity does not match the actual market demand is higher. While the capacity of a factory represents a long-term investment decision, the spread of innovations can fail within a short timeframe or occur far below the expected level. For this reason, it can be assumed that insights from innovation research offer a planning advantage in forecasting the production potential. Regarding an increasing number of global product innovations, the evaluation of empirical data by means of suitable models and methods is becoming more and more accurate in order to reflect typical market patterns based on recurring customer behaviour. This paper takes up these trends and proposes an approach how innovation metrics can be included in the capacity dimensioning process of factory projects. For this purpose, the BASS diffusion model is used to realistically map different market scenarios for the required capacity curves.

Keywords

Factory Planning; Innovation; Diffusion; Capacity Dimensioning, Fuel Cell

1. Introduction

The market for fuel cell systems for vehicles in Germany and Europe is currently still relatively small, especially in comparison with North America and Asia [1]. It is generally assumed that this will change in the near future [2]. Indications of this, are the ambitious German national hydrogen strategy [3] and the EU Hydrogen Strategy [4]. Trencher and Edianto show that there is consensus in industry and research that those policy signals are drivers in market development [5]. In order to drive forward the transformation of energy supply towards fewer emissions, fuel cell technology is well suited, especially in areas where it is otherwise difficult to electrify applications. One of these applications are utility vehicles where fuel cell systems are generally better suited than battery electric systems [6]. Hence, within this article, the focus will be on utility vehicles like trucks and buses.

The future market size, and therefore the sales potential of a company, in a timeframe of five to ten years and beyond is subject to great uncertainty. This influences companies' factory planning. In the worst case, factory planning is based on very optimistic estimates, while the technology remains a niche application or is substituted by other technologies. When focusing on the mobility sector one competing technology is the battery electric vehicle (BEV). In comparison to fuel cell electric vehicles (FCEVs) BEVs are some years ahead in terms of technological maturity and cost reductions [5]. It is expected, that this gap in economic competitiveness will close in the coming decade, even for light-duty vehicles [7]. These forecasts are not certain and rely on continuous technical improvement, which, however, cannot be guaranteed.

In 1998, a miscalculation of future demand and corresponding capacity cost Siemens about $150m [8]. When the decision to invest in a chip plant in Tyneside (UK) was made in 1995, Siemens predicted high demand for 16Mb chips and constant prices. Due to a sharp drop in prices (95% between 1995 and 1998); Siemens was forced to close its plant in Tyneside. This misjudgement was due to an incorrect assessment of the cycles of the traditionally fluctuating demand in the semiconductor industry [9]. Such errors are financially very difficult or even impossible to bear, especially for small and medium-sized companies.

Especially factory and production process planning for innovative products, such as fuel cell components, face the challenge of accurate forecasts because historical or empirical market data do not exist. This leads to a higher risk that capacity and demand do not match, which can lead to expensive adjustment measures.

To mitigate the risk of a capacity and demand mismatch, we suggest incorporating insights from innovation research into traditional factory planning methods with particular focus on capacity dimensioning. For this purpose, this paper presents a short introduction to capacity dimensioning methods and diffusion models. Chapter 4 presents a linkage of both domains concerning the factory planning for the assembly of fuel cell stacks. Chapter 5 compares the advantages and disadvantages of the proposed method. The paper concludes with a summary of the previous points and suggests future research activities on the topic.

2. Capacity Dimensioning

Several authors have developed structured overviews on standardized stages and steps of the factory planning process [10,11]. Moreover, the specific planning requirements and contents are formalized and described in the norm VDI 5200 [12] by the German Society of Engineers which also refers to the official scale of fees for services by architects and engineers (HOAI [13]). Based on these documents, the planning and estimation of required capacity can generally be seen as a base for subsequent structure planning. When investigating these aspects in more detail it seems useful to distinguish between the terms capacity dimensioning and capacity planning to avoid confusion. While the main objective of the latter is the optimal allocation of a fixed capacity according to cost criteria, delivery reliability or flexibility [14], the capacity to be dimensioned refers to the strategic orientation of serving a forecast customer demand without gaps and also in compliance with overall corporate objectives [15]. Accordingly, capacity dimensioning can be interpreted as a discipline of factory planning, the subject of which is the balancing between an order quantity (capacity requirement) and the output quantity achievable through production factors (available capacity).

Major disadvantages of these sequential procedures are information losses on the interface between different planning partners who are, for instance, responsible for production infrastructure, HVAC and plumbing. Furthermore, it leads to an additional effort caused by iterative decision processes. As an answer to this, several concepts for integrated factory planning were developed and refined. An exemplary methodological framework is presented by Wiendahl et al. [11] with the so called synergetic approach to factory planning which comprises a two-dimensional project planning concept that relies on early cooperation between construction planning and production planning. Consequentially, the aspired cooperation necessitates a consolidated base for the planning and construction operations. With regard to this, the methodological and software-related framework of Building Information Modeling (BIM) serves as a basis for collaboration. The associated XML-based and openBIM-oriented scheme, named Industry Foundation Classes (IFC), is currently oriented to the development of buildings and their infrastructure. Thus, this may provide the starting point for the conjoint modelling of the production facilities infrastructure as well as logistic elements.

In any case, a flexible and precise forecast of product demands for infrastructure dimensioning is necessary in order to design and build a flexible and future-oriented production facility.

The basis for planning is a quantification of future demand, which can be carried out using various forecasting methods. Depending on the application, Schönsleben [14] suggests a past- or future-oriented approach. The former is used if valid consumption data for the specific product can be accessed. These are transferred into time series and evaluated by means of mathematical or graphical procedures. If, on the other hand, such values are not available with the necessary validity, a forward-looking method can be used. A characteristic feature of these procedures is that available information on future demand trends is recorded and modelled as extensively as possible. In terms of methodology, both mathematical models and intuitive approaches, such as estimation based on empirical values, have become established. It should be emphasized that the resulting forecast data is subject to the limitations of the method used and therefore only partially addresses the complexity and interdisciplinarity of the influences on market demand.

The forecast results are then transferred into a formal production program which, in addition to the quantity data, also contains product-specific, value-based and time-relevant specifications [10]. At the same time, data for determining the available capacity must be prepared. This is limited by the technical performance of the available resources, and if necessary, personnel organization specifications and budget restrictions can also influence the available capacity [16]. With regard to this, it must also be taken into account which manufacturing processes can be considered for individual production steps and whether this will be accompanied by future replacement investments. The collected data constitutes the basis for the concept-planning phase of factory planning. The data is used for dimensioning the capacity and derived space requirements. More concretely, this phase consists of the sub-steps of the technological, temporal and organizational comparison between capacity requirements and available capacity [11]. First, the information recorded in the available capacity is concretized in terms of the production processes that are used and the required operating resources (machines, robots, tools). The choice of technology creates the necessary conditions to harmonize the temporal premises of the production system in the next step. In conjunction with the personnel requirements of the individual workstations, different shift system variants are created and compared based on qualitative as well as quantitative criteria. Taking into account production-reducing factors such as rework or failures, the existing net-working time T_{Mi} is calculated by multiplying the gross operating time T_{MGi} and the time utilization factor η_{Ri} [17]:

$$T_{Mi} = T_{MGi} * \eta_{Ri} \qquad (1)$$

At the same time, the required occupation time on the individual work stations is to be determined in terms of planned volumes linked to product-specific parts lists [11]. As formula 2 shows, this results from the sum of the set-up time T_{Cji} and the total production time T_{Pji} of a period, which depends from the forecasted production volume x_{ji} and the processing time per unit t_{uji}.

$$T_{Ri} = \sum_{j=1}^{J} T_{Cji} + T_{Pji} = \sum_{j=1}^{J} T_{Cji} + \sum_{j=1}^{J} (x_{ji} * t_{uji}) \qquad (2)$$

The quotient of occupancy time T_{Ri} and net-working time T_{Mi} ultimately leads to the number of operating resources or personnel required (n_i).

$$n_i = \frac{T_{Ri}}{T_{Mi}} \qquad (3)$$

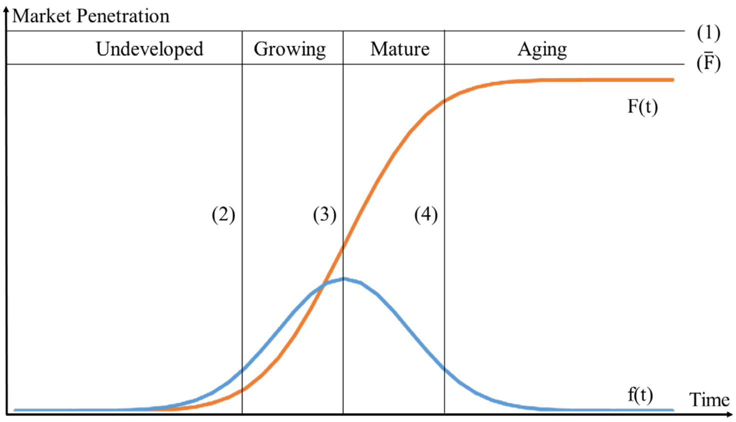
Since the demand forecast on the market does not usually follow an ideal, uniform course, the extent to which capacity should be adaptable in the future must also be taken into account. On the one hand, this concerns the provision of resource-bound flexibility (e.g. short-term changes in the shift model) [18], on the other hand, the strategic positioning in competition is decisive for the capacity expansion of the company.

General alternative strategies for this are *lead* (keeping excess capacity), *match* (demand-synchronous adjustment) and *lag* (delayed, risk-averse adjustment) [19]. The choice of strategy opens up the scope of action to shape the capacity decision depending on the product, the market environment and the company's goals.

The proven planning process shows that the demand forecasts collected at the beginning of the process consistently have a significant influence on the results of the individual steps. Accordingly, the demands for validity and realism are justified. This requirement is particularly challenging for innovative products and young markets. In these cases, it is typically uncertain whether the product will be able to successfully establish in the market and thus achieve market penetration. At the same time, typical competitive situations occur more frequently in young markets, which can strongly change the market distribution in the short term [20]. Since the typical forecasting methods reach the limits of their ability to depict these situations, innovation research has been dealing with the modelling of market developments for a long time. In combination with the capacity dimensioning process described above, these models can also help to consider product-typical demand trends in terms of capacity and to include changes in trends, for example due to specific innovation drivers, in planning at an early stage.

3. Overview of Innovation Diffusion Models

Innovation diffusion can be defined as “[…] the process by which an innovation is communicated through certain channels over time among the members of a social system” [21]. First popularized by Rogers in his 1962 published book *Diffusion of Innovations*, the theory became a staple in economics, social and communication sciences. The concept of innovation diffusion can be used on a micro level to describe the behaviour of individuals, but also on a macro level to describe how an innovation spreads across an entire social group or market. Rogers coined terms for five individual groups of innovation adopters and assumed that their distribution pattern corresponds to a normal distribution. Since then different authors suggested mathematical models to describe those patterns in detail. Figure 1 shows the general components of all diffusion models.



Symbol	Description
(1)	Whole Population e.g. of EU, China. Usually just a fraction of (1) are potential adopters $\bar{F}$
(2)	Lower threshold – if the adoption process reaches this level it will progress till completion
(3)	Infliction point – f (t) reaches its maximum, while the first degree derivative of $\bar{F}$ changes to negative
(4)	Upper threshold – adoption process is almost complete the rest of the adoption process cannot be exactly modelled
$\bar{F}$	Potential adopter population
F (t)	Cumulative adopter function
f (t)	Non-cumulative adopter function (diffusion function)

Figure 1: General components of diffusion models, based on [22]

Diffusion models are used to predict the sales of new technologies. Other methods, such as trend extrapolation, cannot be used because empirical values are not available. The use of diffusion models is thus classified as a future-oriented method for forecasting.

The project was intended to provide an initial proof of work of integrating diffusion models into capacity dimensioning. Therefore, to facilitate the validation of the results, only simple diffusion models were considered. These are usually based on the logistic distribution or derivatives thereof. Three of these models and their advantages and disadvantages will be briefly presented here.

- **Bass**
  Based on the work of Rogers, Bass extended the general model of different innovation stages with a mathematical foundation [23]. The Bass diffusion model turned out to be highly influential and is used in various industries to forecast future sales. This model replicates empirically proven progressions of the innovation course of products sufficiently accurately. Bass tested this himself using sales figures for 11 consumer products [23]. This model needs three different input parameters. Those are the coefficient of innovation (p), the coefficient of imitation (q) and the population of adopters (m). With sufficient data for the values p and q (e.g. from adoption processes of similar products), the model is generally very accurate. The model is less well tested for industrial goods [24]. However, see point 4 for an overview of comparable publications in which the Bass model was used to forecast similar technologies.
- **Fisher-Pry**
  The Fisher-Pry model makes three assumptions. First, it assumes that one technology is substituted by another to satisfy the same need. Second, if the substitution process reaches a certain threshold, then the process continues until full substitution occurs. Third, the substitution rate is proportional to the remaining quantity of the old product compared to the new one [25]. The model is characterized by its ease of use. To utilize it, only two input parameters need to be determined. One is the growth rate at the beginning of the diffusion process and the other is an estimate of the year in which the market penetration will be half [24]. However, compared to the Bass or Blackman-Mansfield models, the model yields less meaningful results [24].
- **Blackman-Mansfield**
  In itself, Blackman's model does not represent a diffusion model. In contrast to the models mentioned so far, it does not examine the substitution or adoption of a product, but the development of the performance of a technology, expressed in a figure of merit [26]. However, Blackman, who bases his model on the work of Mansfield [27], himself shows the connection between technical progress and market substitution [26]. Compared to the Bass model, the factor of cost between different technology options is considered, which can be an advantage [24]. However, the model produces overly optimistic values at the beginning of the prediction and overly pessimistic values at the end [22].

As indicated earlier, the models shown here represent fundamental work in the field of product diffusion. These models have the crucial disadvantage that they are based on a fixed mathematical form. Various authors have tried to compensate for these disadvantages by modifying individual models, by combining different models or by creating completely new models. For an overview, see e.g. [22]. For this paper, these models were not considered.

For the prototypical application, the Bass model was chosen. This is because there are a number of publications that model comparable technologies with its help (listed in table 2).

The values determined with the help of these methods are of course already used in production program planning and thus influence capacity dimensioning. However, normally only the static data determined once is used. For the factory planner, this results in a rigid numerical framework, although changes in market adoption can occur very quickly, especially with new technologies. By integrating diffusion models into the planning process, more differentiated statements can be made, especially with regard to strategies for capacity flexibilization, since this makes it possible to include different adoption processes during planning.

The following section describes the integration of the bass diffusion model in the capacity planning for an assembly line of fuel cell stacks.

4. Application in the assembly of fuel cell stacks

Although the fuel cell was invented as early as 1839 by William Grove [28] and has since been tested many times in various fields of application, its use for a broad market has only recently been pushed. The driving force behind this development can be identified primarily in a growing social awareness of the environment, which is leading to an increasingly critical questioning of conventional drive technologies. At the same time, today's technologies and findings enhance the performance of the fuel cell and can limit application-related dangers and disadvantages. For these reasons, many indicators argue for a disruptive innovation character [29], which exhibits promising future potential in mobile, portable and stationary applications.

At the same time, the technology has disadvantages that could impair rapid diffusion from today's perspective. On the one hand, the production of fuel cell drives is very cost-intensive compared to other drive concepts. Cost drivers are the raw materials needed to manufacture the multi-electron unit [30], the pressure-resistant tank system [31] and the environmentally friendly production of the hydrogen [32]. On the other hand, historical market data show that gaseous fuels tend to be avoided because of the hazards associated with them [33]. Factors such as these endanger adoption in the passenger car market. However, the implications of these factors are much less significant for the operation of fuel cells in trucks or buses. According to a study, which was commissioned by the German state of Baden-Württemberg, this market is growing faster than the passenger car market [34]. Based on the study, two scenarios were defined, which show the course of demand for a pessimistic and optimistic development. Since the characteristics of these scenarios are relatively static and innovation-typical influences are largely neglected, the values were inserted into the model equation according to BASS to determine the model parameters for innovation (p), imitation (q) and potential market size (m). A geometric average was taken over all ten generated value triplets. As a result, the following values can be determined:

Table 1: Diffusion parameters derived through regression analysis

Scenario	p	q	m
Minimum	0,016164	0,522596	100.000
Maximum	0,025612	0,547746	250.000

As the plausibility of the values can only be derived from the numerical amount to a limited extent, publications with similar applications were used as a basis for comparison. Table 2 gives an overview of the authors with the corresponding reference product. Although there is a high degree of scattering among the publications, it can be stated that the values determined fit the model applications typical for the industry.

Table 2: Diffusion parameters for similar products

Authors	Product	p	q	m
Lukas et al. [35]	Electric vehicle batteries	0,022	0,413	2.150.00
Massiani & Gosh [36]	LPG-vehicles in Germany	0,0779	0,3718	75,566
McManus & Senter [37]	Plug-in hybrid electric vehicle	0,00262	0,70935	1.922.806
Li, Chen & Zhang [38]	BEVs in China	0,0013	0,0839	5.000.000
Becker et al. [39]	BEVs	0,025	0,4	2 scenarios

Using the determined model variables, sales figures can be forecasted according to the expected diffusion course and included in the production program. In this specific example, the production program will focus in particular on the manufacturing of the fuel cell stack (FCS). As one of the main components of the fuel cell drive, an FCS represents a combination of single fuel cells and is therefore composed of a defined number of membrane electrode assemblies (MEA), bi-polar plates, seals and end plates [40]. The various components are pressed together in a multi-stage process and then subjected to a leakage test [41]. As an example, the stacking system will now be used to show how the change of diffusion parameters influences the capacity of the production system to be planned. Assuming a 2-shift system and a process time of 30 min per FCS, this can be determined by the number of systems required (Figure 2).

Parameters	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Year
p	0,016164	0,025612	0,038418 (+50%)	0,025612	0,025612	
q	0,522596	0,547746	0,547746	0,821619 (+50%)	0,547746	
m	100.000	250.000	250.000	250.000	375.000 (+50%)	
Number of Stacking Systems	1	1	1	1	1	1
	1	1	1	1	1	2
	1	1	2	2	2	3
	1	2	3	3	3	4
	1	3	4	4	4	5
	1	4	5	6	6	6
	2	5	6	7	7	7
	2	6	7	8	9	8
	3	7	8	8	10	9
	3	8	8	8	11	10

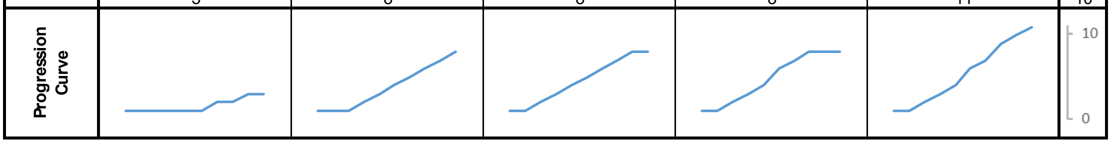


Figure 2: Correlation between diffusion parameters and needed capacity of FCS stacking systems

5. Discussion

Figure 2 shows five different scenarios with their parameters (p, q, and m), the number of required stacking systems and the progression curve for a better visualization of the development of the required capacity. Scenario 1 corresponds with the minimum scenario in table 1. In the first ten years, the required capacity is slowly increasing only in the seventh year an additional system is required. In this scenario, it is not advisable to invest in a second system from the beginning. Especially considering that the price for one system is about 700.000 € [42] leading to high and unnecessary capital commitment costs. Scenario 2 is in line with the maximum scenario from table 1. Here, compared to scenario one, a significantly faster increase in the number of systems required can be seen, with one system still being sufficient in the first three years. Scenarios 3 to 5 show different variations of scenario 2. The parameters of p, q and m are successively increased by 50%. The change of the innovation parameter p leads to the smallest impact on capacity dimensioning. Increasing parameter q in scenario 4 results in a similar outcome compared to scenario 3. Scenario 5 shows the fastest increase in required systems, showing that the market size is the most important factor. Therefore, special attention should be paid here to the determination of this parameter.

The advantage of the presented method compared to conventional capacity dimensioning is that by varying the parameters, more differentiated statements on the capacity requirements of operating resources become possible. This represents a considerable improvement over statements based only on static sales forecasts. The method is a relatively simple procedure that can be applied in small and medium-sized companies. Typically, such companies do not have a large market research department, which means that they rely on freely available or paid studies when planning the sales of innovative products. With the help of the presented method, the data from these studies can be processed and used to compare different possible diffusion patterns and thus capacity curves.

Despite the fact that the capacity dimensioning has been made more flexible by the method presented above, the consideration is still very static. The influence of other variables, such as politically desired expansion targets for fuel cell fleets, can or even must be considered. Despite the relative simplicity, expert knowledge is rather necessary compared to the use of ready-made studies, which leads to a greater effort. More complex statements and progressions can only be achieved by using other diffusion models. However, these require a more in-depth knowledge of diffusion research methods.

6. Conclusion

The importance of alternative drive technologies is increasing. Currently battery-electric vehicles are in the spotlight of companies, research and buyers. Fuel cell technology is another contender to make mobility more sustainable. Companies that want to invest in this market are faced with the challenge of assessing how the market will develop. Market developments also determine how much production capacity is needed. Working with static figures alone for dimensioning can lead to misjudgements and thus to malinvestments. To avoid this, a general method was presented in the paper that integrates insights from innovation management into factory planning. For that purpose, the paper presented the traditional approach for capacity dimensioning in factory planning projects. A short overview was given on the subject of innovation diffusion models, as a group of models suitable for forecasting future demand of innovative technologies. Various scenarios were set up using the Bass Diffusion model to size the demand for stacking systems in the assembly of fuel cell stacks. The values for this were determined using a regression analysis from a public study. The plausibility of the regression analysis was checked by comparing the values with those of other similar technologies. The method results in different curves for the need for stacking systems over time. Based on the curves, factory planners can make a better estimate of the flexibility strategy to be selected.

Future research and application of this method should concentrate on following aspects. Currently, only variations of the parameters of the Bass diffusion model are considered. They can be seen as a summary of different factors and influences, but a further differentiation of the influencing factors can be made. For further modelling, a system dynamics approach can be used to investigate the influencing factors, which may also result from legal and political conditions. Currently, the described method exists only as an application in Excel. However, in order to enable the dissemination in practice and to improve the usability, it is advantageous to implement the method in existing software solutions. The extent to which the method can be integrated into already established factory planning processes requires further investigation. For this purpose, the use in a real application scenario is desirable.

Acknowledgements

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

References

[1] E4tech, 2019. The Fuel Cell Industry Review 2019.

[2] Saritas, O., Meissner, D., Sokolov, A., 2019. A Transition Management Roadmap for Fuel Cell Electric Vehicles (FCEVs). J Knowl Econ 10 (3), 1183–1203.

[3] Federal Ministry for Economic Affairs and Energy, 2020. The National Hydrogen Strategy. Federal Ministry for Economic Affairs and Energy. https://bit.ly/2QKNd8B. Accessed 2 March 2021.

[4] European Commission, 2020. A hydrogen strategy for a climate-neutral Europe. European Commission. https://bit.ly/3m3pNXA. Accessed 2 March 2021.

[5] Trencher, G., Edianto, A., 2021. Drivers and Barriers to the Adoption of Fuel Cell Passenger Vehicles and Buses in Germany. Energies 14 (4), 833.

[6] Ajanovic, A., Haas, R., 2021. Prospects and impediments for hydrogen and fuel cell vehicles in the transport sector. International Journal of Hydrogen Energy 46 (16), 10049–10058.

[7] Morrison, G., Stevens, J., Joseck, F., 2018. Relative economic competitiveness of light-duty battery electric and fuel cell electric vehicles. Transportation Research Part C: Emerging Technologies 87, 183–196.

[8] Charles, D., Benneworth, P., 1999. Plant Closure and Institutional Modernisation. Local Economy 14 (3), 200–213.

[9] Loewendahl, H.B., 2001. Bargaining with Multinationals. Palgrave Macmillan UK, London.

[10] Schenk, M., Wirth, S., Müller, E., 2010. Factory Planning Manual. Springer Berlin Heidelberg, Berlin, Heidelberg.

[11] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2015. Handbook Factory Planning and Design. Springer Berlin Heidelberg, Berlin, Heidelberg.

[12] Verein deutscher Ingenieure e.V., 2011. VDI 5200: Fabrikplanung - Planungsvorgehen. VDI-Verlag, Düsseldorf.

[13] Bundesamt für Justiz, 2013. Honorarordnung für Architekten und Ingenieure: HOAI.

[14] Schönsleben, P., 2016. Integral Logistics Management. CRC Press, Boca Raton.

[15] Abramowski, M., 1998. Demand oriented factory planning through reference models, in: Schönsleben, P., Büchel, A. (Eds.), Organizing the Extended Enterprise. Springer US, Boston, MA, pp. 212–221.

[16] Holtewert, P., Bauernhansl, T., 2016. Increase of Capacity Flexibility in Manufacturing Systems by Substitution of Product Functions. Procedia CIRP 57, 92–97.

[17] Grundig, C.-G., 2018. Fabrikplanung: Planungssystematik - Methoden - Anwendungen, 6.th ed. Hanser, München, 333 pp.

[18] Zaeh, M.F., Moeller, N., Muessig, B., Rimpau, C. Life cycle oriented valuation of manufacturing flexibility, in: 13th Cirp International Conference, pp. 699–704.

[19] Gong, Y., 2013. Global Operations Strategy. Springer Berlin Heidelberg, Berlin, Heidelberg.

[20] Brüggemann, J., Meub, L., 2017. Experimental evidence on the effects of innovation contests. Information Economics and Policy 39, 72–83.

[21] Rogers, E.M., 2003. Diffusion of innovations, 5th ed. Free Press, Riverside.

[22] Jaakkola, H., 1996. Comparison and Analysis of Diffusion Models, in: Kautz, K., Pries-Heje, J. (Eds.), Diffusion and Adoption of Information Technology. Springer US, Boston, MA, pp. 65–82.

[23] Bass, F.M., 2004. A New Product Growth for Model Consumer Durables. Management Science 50 (12_supplement), 1825–1832.

[24] Gilshannon, S.T., Brown, D.R., 1996. Review of methods for forecasting the market penetration of new technologies. Pacific Northwest National Laboratory, Richland. https://bit.ly/3u5eQHV.

[25] Fisher, J.C., Pry, R.H., 1971. A simple substitution model of technological change. Technological Forecasting and Social Change 3, 75–88.

[26] Blackman, A.W., 1971. A mathematical model for trend forecasts. Technological Forecasting and Social Change 3, 441–452.

[27] Mansfield, E., 1961. Technical Change and the Rate of Imitation. Econometrica 29 (4), 741.

[28] Grove, W.R., 1838. LVI. On a new voltaic combination. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 13 (84), 430–431.

[29] Hardman, S., Steinberger-Wilckens, R., van der Horst, D., 2013. Disruptive innovations: The case for hydrogen fuel cells and battery electric vehicles. International Journal of Hydrogen Energy 38 (35), 15438–15451.

[30] U.S. Department of Energy (DOE), 2020. 2019 Annual Progress Report: DOE Hydrogen and Fuel Cells Program. U.S. Department of Energy (DOE). https://www.nrel.gov/docs/fy20osti/76055.pdf.

[31] Chanaron, J.-J., 2020. Challenges to radical innovation: A car for the future: Part one. Megatrend revija 17 (2), 1–30.

[32] Hydrogen Council, 2020. Path to Hydrogen Competitiveness. A Cost Perspective. Hydrogen Council. https://www.h2knowledgecentre.com/content/policypaper1202.

[33] Gnann, T., Wietschel, M., Kühn, A., Thielmann, A., Sauer, A., Plötz, P., Moll, C., Stütz, S., Schellert, M., Rüdiger, D., Waßmuth, V., Paufler-Mann, D., 2017. „Brennstoffzellen-Lkw: kritische Entwicklungshemmnisse, Forschungsbedarf und Marktpotential“. https://bit.ly/3m5vynC.

[34] Weichenhain, U., Lange, S., Koolen, J., Benz, A., Hartmann, S., Heilert, D., Henninger, S., Kallenbach, T., 2020. Potenziale der Wasserstoff- und Brennstoffzellen-Industrie in Baden-Württemberg.

[35] Lukas, E., Spengler, T.S., Kupfer, S., Kieckhäfer, K., 2017. When and how much to invest? Investment and capacity choice under product life cycle uncertainty. European Journal of Operational Research 260 (3), 1105–1114.

[36] Massiani, J., Gohs, A., 2015. The choice of Bass model coefficients to forecast diffusion for innovative products: An empirical investigation for new automotive technologies. Research in Transportation Economics 50, 17–28.

[37] McManus, W., Senter, R., Jr., 2009. Market Models for Predicting PHEV Adoption and Diffusion. https://bit.ly/3ru6ojO. Accessed 2 March 2021.

[38] Li, S., Chen, H., Zhang, G., 2017. Comparison of the Short-Term Forecasting Accuracy on Battery Electric Vehicle between Modified Bass and Lotka-Volterra Model: A Case Study of China. Journal of Advanced Transportation 2017, 1–6.

[39] Becker, T.A., Sidhu, I., Tenderich, B., 2009. Electric Vehicles in the United States: A New Model with Forecasts to 2030. https://bit.ly/39qEZcd. Accessed 2 March 2021.

[40] Lee, S.-J., Hsu, C.-D., Huang, C.-H., 2005. Analyses of the fuel cell stack assembly pressure. Journal of Power Sources 145 (2), 353–361.

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

[42] James, B.D., Huya-Kouadio, J.M., Houchins, C., Desantis, D.A., 2017. Mass Production Cost Estimation of Direct H2 PEM Fuel Cell Systems for Transportation Applications: 2017 Update.

Biography

Maximilian Stange (*1992) studied industrial engineering in Zwickau and at the Universitat de València, followed by a master's degree in logistics at the Otto von Guericke University in Magdeburg. Since 2019, he is working as a Research Associate at Fraunhofer IWU. In his research, he focusses on economic and ecologic assessment of production systems and novel business models.

Lukas Roth (*1995) completed a bachelor's and master's degree in industrial engineering at TU Chemnitz. Since 2018, he is working as a student assistant at Fraunhofer IWU. As part of this work, he recently completed his master's thesis in the field of "Capacity Dimensioning under Forecast Uncertainty" and is also researching factory planning issues related to sustainability and digitalization.

Marian Süße (*1992) started working at the Fraunhofer-Institute for Machine Tools and Forming Technology IWU as a Student Assistant in 2014. Since 2017 he is working as Research Associate in the Department Factory of the Future where he is in charge of the Group for Factory Design and Simulation since 2018. In his research and PhD-Thesis, he focusses on the development of methods for factory layout planning and layout optimization approaches as well as ecologic impacts of production systems.

Steffen Ihlenfeldt (*1971) studied at the TU Braunschweig and has been a member of the Fraunhofer Institute for Machine Tools and Forming Technology IWU in various positions since 1997 at. He received his Doctorate in 2012. Since 2015 Prof. Dr.-Ing. Steffen Ihlenfeldt has been called to the Position of Professorship for Machine Tool Development and Adaptive Control systems at the TU Dresden. In addition to his work at the TU Dresden, he is Institute Director of the Fraunhofer IWU, responsible for the scientific field of Production Systems and Factory Automation.

2nd Conference on Production Systems and Logistics

Electrostatic Self-Assembly Technique for Parallel Precision Alignment of Optical Devices

Martin Stucki[1,2], Christoph Schumann[1,2], Annika Raatz[1,2]

[1]Leibniz University Hannover, Institute of Assembly Technology, An der Universität 2, 30823 Garbsen, Germany
[2]Cluster of Excellence PhoenixD (Photonics, Optics, and Engineering – Innovation Across Disciplines), Hannover, Germany

Abstract

In precision assembly, the cost of machine technology increases significantly when high assembly accuracy is required (<15 µm). One reason is that higher accuracy with conventional automation technology requires much more precise and expensive machine components, such as bearings and actuators. Electrostatic self-assembly is a technique for the automatic alignment of micro-components without the need for precise machines and thus has the potential to reduce fabrication costs significantly. With this technique, electrodes are placed on the micro-components and the substrate. A low viscosity fluid is applied to the substrate and the components are roughly positioned. One pair of electrodes on the component faces one pair of electrodes on the substrate, equivalent to plate capacitors connected in series. If an alternating voltage is applied to the substrate electrodes, an electric field is formed. This results in electrostatic attraction in the transversal and lateral direction, which leads to an alignment of the components on the substrate. In this paper, we describe the structure design process for electrostatic self-assembly. Instead of micro-components, we use a rectangular glass wafer with a length of 125 mm. Within two test series, we prove that the existing technique is also suitable for a larger scale.

Keywords

Precision Alignment; Self-Assembly; Parallel Assembly

1. Introduction

Optical systems, in combination with electronic circuits, are becoming increasingly important. Growth markets are camera technology for smartphones and autonomous driving or sensors for consumer electronics. More than twenty percent of US citizens are currently using wearables, mostly with integrated optical heart rate sensors for health monitoring [1]. This is just one example of a rapidly spreading technology in recent years, and several research groups are currently working on the development of new optical systems in the fields of sensor technology, signal processing and health care [2–4]. Efficient production technology is important for further mainstreaming these new technologies. The production of semiconductor materials and silicon are incompatible with the manufacturing processes of many optical devices. For example, the 3D shape of almost all lenses conflicts with the pure surface technology of the LIGA process. For this reason, separate processes are necessary for assembly and connection [5].

Sequential assembly and positioning in the micrometre range with the aid of positioning devices is very time-consuming and the machine technology expensive [6]. For this reason, several years ago, research groups researched parallel assembly of micro parts without direct handling and published under the term

DOI: https://doi.org/10.15488/11255

electrostatic self-assembly. Primarily, fluidic and electrostatic driven mechanics were in the focus of the experimental research for microelectromechanical systems (MEMS) [6] and DALIN et al. build a simulation, based on experimental data [7,8]. Within the Cluster of Excellence PhoenixD, the Institute of Assembly Technology develops handling processes for optical devices and micro parts. Within a new approach, we adapt the electrostatic self-assembly concept for aligning whole pre-processed wafers instead of the single handling of small parts like MEMS or chips. Due to the novelty of this approach, the focus of this paper is to present the concept for the full-wafer alignment, the development of a structure design and to prove the suitability of the process for a different scale.

2. Concept of electrostatic Self-Assembly

The term self-assembly is mainly used in chemistry, e.g. when nanoparticles selectively localize within a particular microdomain, to form patterns or nano-structured coatings on a surface [9,10]. In the engineering or production context, self-assembly is a method to align components without direct handling. In this paper, we address electrostatic self-assembly as a system consisting of a substrate (bottom wafer) and a component (top wafer) separated by a fluid. This fluid serves several functions: it reduces friction and thus enables movement in the lateral direction. In addition, it acts as a dielectric for the electric field [6]. On both wafers, conductive structures are deposited. These structures consist of conductive tracks and surfaces called pads. Two facing pads form a plate capacitor [8,11]. If a pad is electrically polarised, e.g. negatively charged, an opposite charge is attracted to the facing pad due to the Coulomb force. This leads to a locally induced charge shift on the secondary conductive structure. This means the transfer of electric potential between the pads is capacitive and therefore only one side needs a galvanic connection to the power supply. Due to the design of the plate capacitor, an electrostatic attraction takes place between the pads. As a result, forces act in the lateral and transverse directions. The component starts to moves within an electric potential field towards the energetically most favourable state (Figure 1). We artificially create the electrostatic field and define the position with the lowest potential as the assembly position with the structure design [12,13]. Once the assembly position is reached, the component is trapped within an equilibrium of the interacting forces.

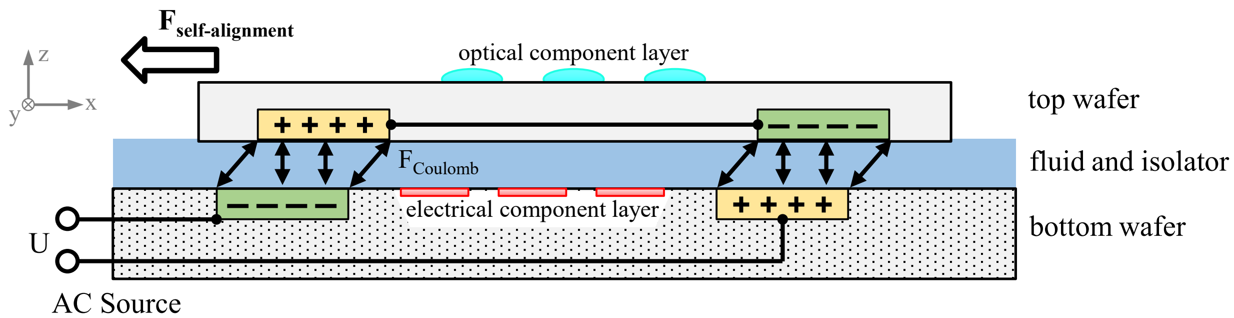


Figure 1: Concept for self-assembly of parts with electrostatic attraction

3. Full-wafer approach and structure design

Electrical components, such as CMOS sensors, are manufactured on wafers with a high degree of precision. A similar process is possible for optical components on glass wafers. Instead of assembling these electrical and optical components individually with a high-precision pick-and-place process, it is more economical to perform parallel alignment and bonding. The possibility of aligning small chips with electrostatic self-assembly has already been demonstrated [8]. Nevertheless, the parallel assembly of already diced microsystems still requires a sequential pre-positioning of the individual parts. To bypass this time-consuming placing process, this article took a novel approach to achieve the alignment of an entire wafer.

Figure 2 shows the concept of this approach. In the centre of the wafer, functional component layers can be applied using standardized micro-processing techniques. Electrically conductive structures are deposited outside or between the components for the alignment process. When aligning the upper wafer via electrostatic self-assembly, the two component layers are precisely positioned on top of each other. Then the two wafers are bonded together over their entire surface. In the following cutting process, the two bonded wafers can be separated into individual modules and used for further applications.

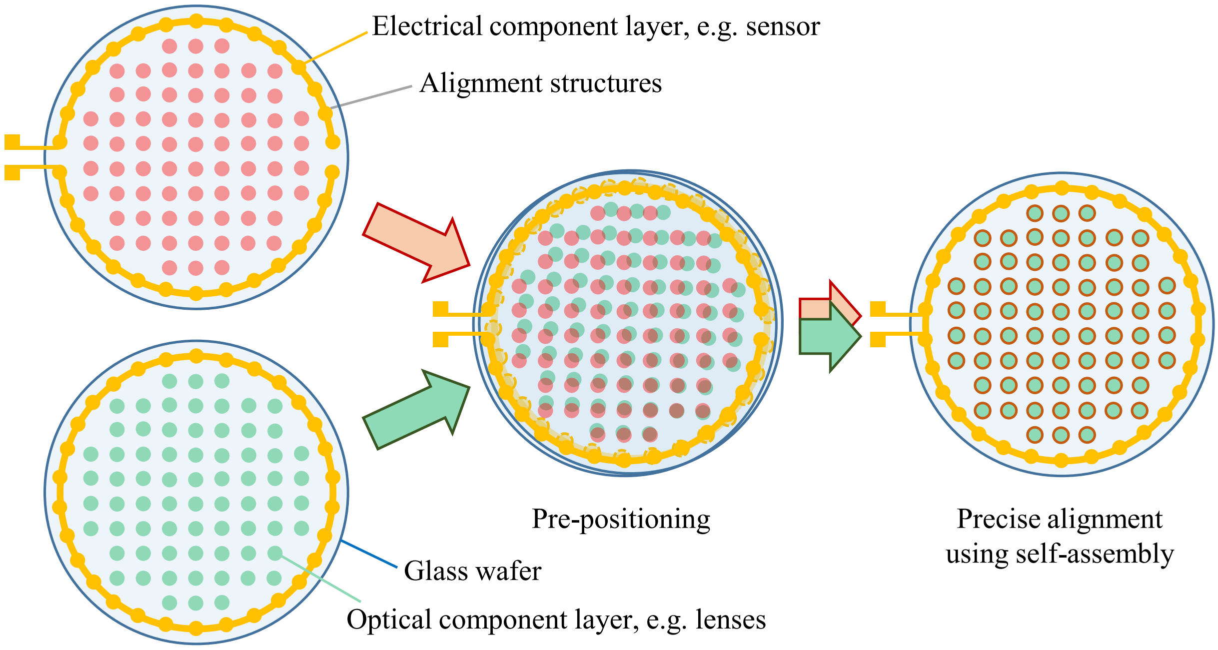


Figure 2: Novel approach for full-wafer alignment via self-assembly

3.1 Systematic structure design

The first step in building a self-assembly system is the structure design. This involves determining the geometry and arrangement of the conductive structures. Although the individual sub-steps have already been applied several times [7,8,11,6], a systematic process description has not been established yet. For this reason, we present a systemic development process. This process includes three steps, which are illustrated in Figure 3:

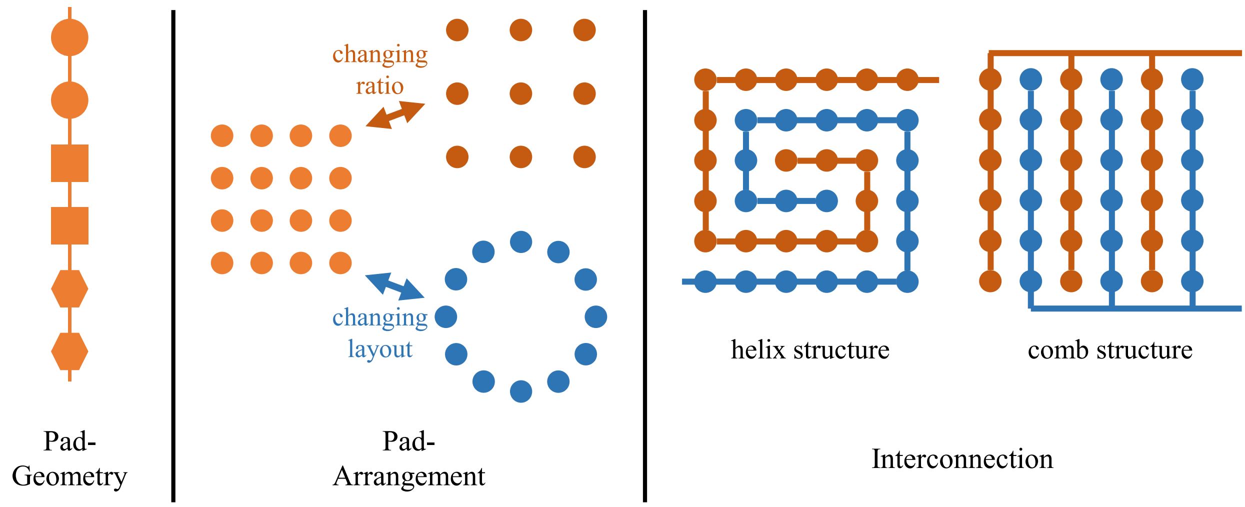


Figure 3: The three steps of structure design

Pad-Geometry:
Pad geometry describes the geometric shapes of the individual opposing active surfaces. These surfaces were build up as squares, rectangles, circles or hexagons [14,11]. In principle, however, any geometry is possible.

Pad-Arrangement:
Pad arrangements are about the spatial positioning of the individual pads on the surface. In previous work, the only arrangement found is in a square [6]. The square shape is derived from the geometry of silicon chips. In addition, in the arrangement step, the pad distance and pad size are defined. These parameters are relevant because two closed by pads can influence each other's electric field. TONDORF and WILDE [6] have taken the ratio of these two values as the decisive parameter.

Interconnection:
Even if the arrangement is already fixed, the pads can be connected in different ways. For electrostatic self-assembly, at least two adjacent pads must have different potentials in at least one direction. For example, in a square arrangement, parallel comb structures or helix structures are possible [6].

3.2 Structure design for full-wafer alignment

Our novel approach for full wafer alignment requires a significant adaptation of the structure design used for the alignment of micro parts. Firstly, due to the significantly larger surface area, the geometry of the structures must be resized to ensure that there is enough active surface to achieve an alignment. Secondly, wafers are round due to the production process, and therefore, the arrangement of individual pads must also be changed to create the maximum possible usable area for the application. We chose a rectangular pad geometry for the experiments (1 x 5 mm²) and arranged 76 pads in a circle with a diameter of 105 mm to leave space in the centre for electrical and optical components in the planned application. The pad distance is on average 3 mm (ratio 1:3) and was specified based on [6]. The pads are connected alternately. In this way, two adjacent pads always have a different electrical potential. We positioned measurement structures in the centre of the wafer. Conductive tracks link the pads on the top wafer. Figure 4 shows the developed structure design.

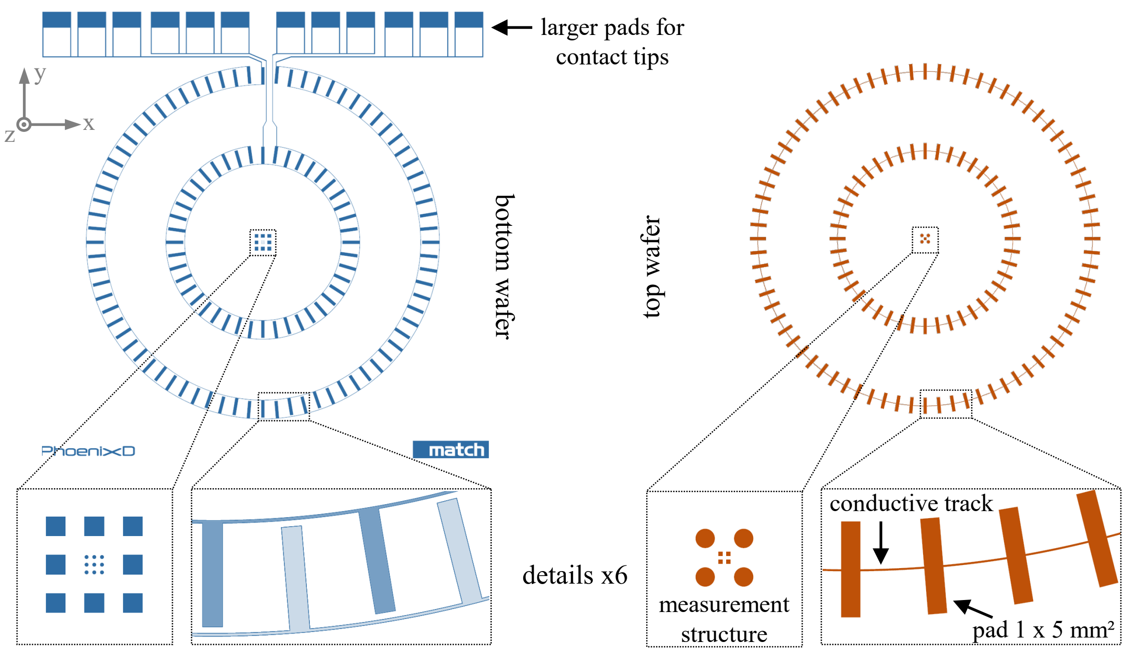


Figure 4: Structure design for full-wafer alignment

The conductive structures were deposited on rectangular exposure masks. The bottom wafer has an edge length of 15.25 cm (~6 inches). The top wafer has an edge length of 12.7 cm (~5 inches) and a thickness of 560 µm. The material in both cases is soda-lime glass with a chromium layer of 0.1 µm, with a thin oxide layer on the surface. On the top side of the bottom wafer, there are large rectangular areas for galvanic contact with contact tips. There is a second circular arrangement of pads with a smaller diameter, but they are not used in the experiments.

4. Experimental setup

We build the experimental setup on a damping plate to reduce mechanical vibrations that would affect the measurement. The bottom wafer lies on a backlight and is illuminated from below. A camera (IDS 3880 CP-6.4 MP) with a 4x telecentric zoom lens and additional coaxial illumination is mounted in the middle above the measurement structures. We can observe the relative position via a monitor in situ. A confocal distance sensor (microEpsilon IFS2405-3) measures the relative distance of the surface below. Two contact tips connect the structures on the bottom wafer to the voltage source. A function generator (GW-Insteck ASR 2050) directly supplies the voltage up to 350 VAC. For higher voltages, a transformer is interposed.

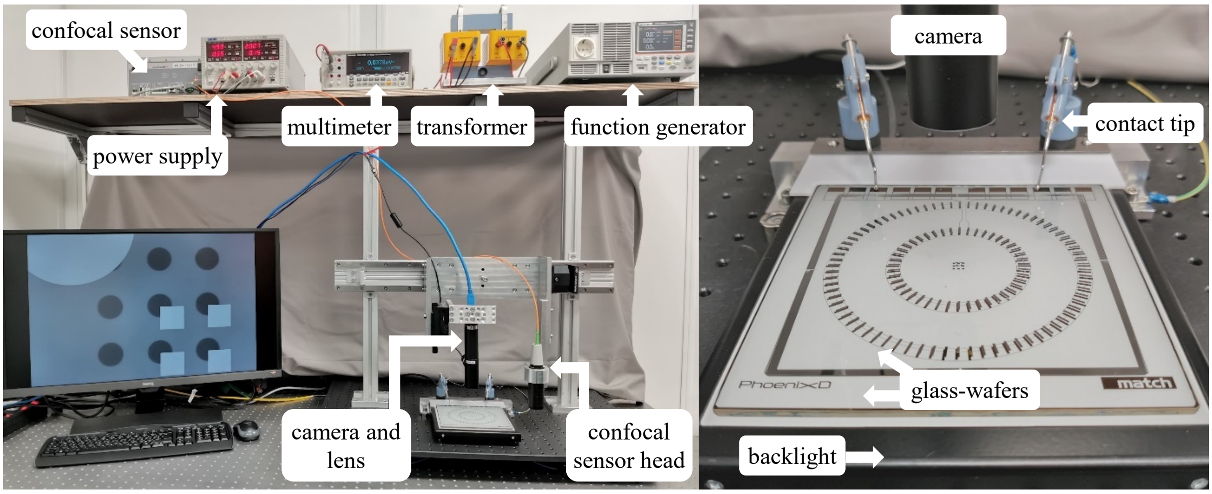


Figure 5: Experimental setup: overview (left) – detail (right)

4.1 Experimental procedure

Two larger test series were carried out for this paper. The first series deals with the influence of voltage, the second with the frequency of the alternating current.

First, the fluid is applied between the substrate and the component. In our experiments, we use the acrylic-based UV adhesive Sicuwell 7043-N, which can be activated by UV light. The curing process is not part of our test. The thickness of the adhesive layer varies between 70 µm and 100 µm in the experiments. Manual pre-positioning takes place before the electrostatic alignment. The initial positioning error is between 450 µm and 700 µm in *x*- and *y*-direction. After activating the voltage, the movement of the top wafer starts without any measurable delay. When the component has reached the final position, the voltage is deactivated. The camera documents the movement during the entire alignment process focusing on the measuring structure in the centre of the wafer alignment, the top wafer is pre-positioned again, and the tests are repeated 10 times per parameter set. The camera images allow measurement of the top wafer's relative position during the entire process.

Combined, the camera with the telecentric lens system has a resolution of 0.6 µm/px, which was confirmed by a separate measurement. The evaluation is done automatically via the measurement structures using the Open Source Computer Vision Library (OpenCV). Within OpenCV, we used algorithms for greyscale analysis to detect the structures and enhance the measuring system's overall resolution by using sub-pixel interpolation and mean values.

4.2 Evaluation criteria for the process

The assessment of the results can be done based on several criteria. The first criterion is the success rate of the positioning. The success rate quantifies the rate of a successful alignment, independent of the achieved accuracy. The most important criterion for our research is the absolute positioning error *d*. It is defined as the absolute distance between the target position and the achieved alignment position. It is calculated as the vector length of the positioning errors in *x*-direction e_x and *y*-direction e_y.

5. Results and Discussion

In both test series – voltage and frequency – we evaluate the effect of the input parameter on the positioning error. Figure 6 shows a typical measurement curve. Placed in the diagram are the position error e_x and e_y, the length of the resulting vector *d* (absolute position error) and the velocity *v* in µm/s, as median over 10 values.

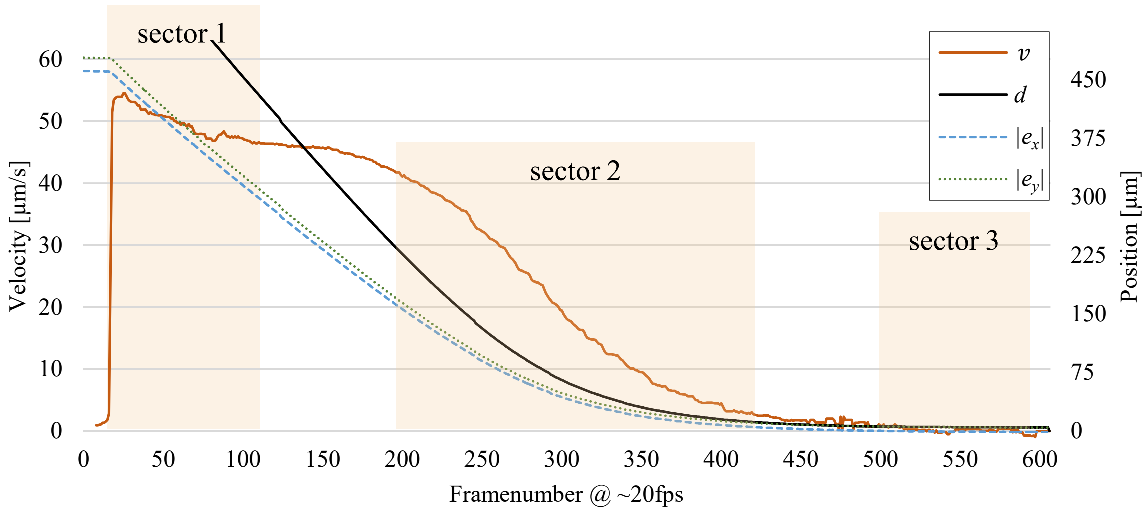


Figure 6: Diagram of the positioning errors e_x, e_y, *d* and the velocity *v*

We classified the measurement into several sectors:

Sector 1 (start): at the beginning, the voltage is activated, and the velocity increases rapidly. The movement in *x* and *y* is almost linear, which corresponds to an almost constant velocity.

Sector 2 (middle): as the top wafer approaches the assembly position, the velocity reduces. The position ascends asymptotically to the final value.

Sector 3 (end): at the end of the measurement, there is no measurable motion. The measured values only fluctuate within the range of measurement accuracy. At this point, the position error is calculated as an average of 10 measured values.

This observation fits into the theoretical model developed by DALIN et al. [14,7,8]. There is only a small overlap of the pads initially, but the pads are still in the electric field range. During the alignment, the overlap of the pads increases. According to the model by DALIN et al., the lateral alignment force decreases with the overlap, i.e. the force decreases during the alignment. At the assembly position, the overlap is 100%, and the force becomes zero. This description corresponds with the measurement curve. The velocity decreases within the process, and movement stops when the assembly passion is reached.

5.1 Influence of the voltage level

The first series of experiments investigates the influence of the voltage level between 250 VAC and 1250 VAC (sine wave, 50 Hz). Figure 7 shows the correlation between the positioning error and the applied voltage.

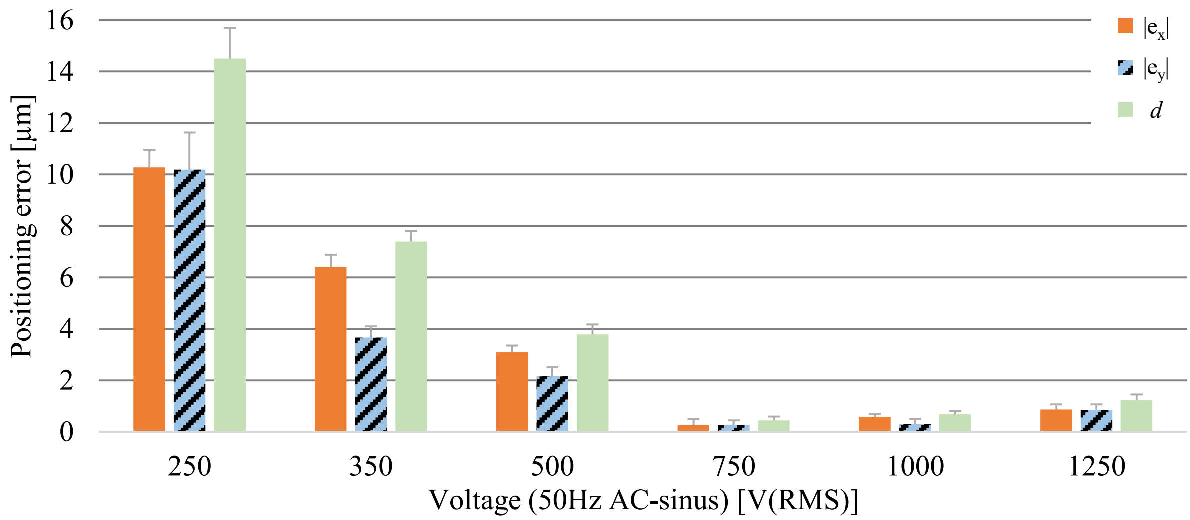


Figure 7: Correlation of the positioning error on the AC-voltage

In this series, the success rate is 100%. This value must be set in relation to the high deviation of up to 700 µm in the pre-positioning. Nevertheless, the top wafer was successfully aligned in every case.

Between 250 VAC and 750 VAC, the positioning error decreases significantly. With 750 VAC, the positioning error *d* is 0.4 µm with a standard deviation of 0.2 µm. However, the measuring method limits the measurement resolution of the positioning error.

At a voltage of over 750 VAC, the positioning error increases again. This indicates that above 750 VAC a different equilibrium of forces is established. With the voltage, the transverse force also increases, i.e. in the *z*-direction. As a result, the upper wafer is attracted and the adhesive layer between the glass wafers is reduced, which leads to an increase in the frictional force. Since an in situ measurement of the adhesive layer's thickness is not possible in the given experimental setup, further investigation is necessary to assess this relationship.

5.2 Influence of the frequency

The second series of experiments investigates the influence of the frequency. With the best results at 750 VAC, the positioning error has reached the limit of our measuring method. To detect a change in the positioning error caused by the frequency, we had to decrease the voltage. Additionally, up to 350 VAC we can use the function generator directly within the setup and the waveform and frequency is not influenced by the transformer. Therefore, the voltage is set to a constant sinus wave at 350 VAC.

The experiments are done in a range of 2 Hz to 200 Hz. Figure 8 shows the correlation between the positioning accuracy and the frequency. As in the first series, the success rate is 100%. The results show that the frequency influences the positioning error. The best result is achieved with 25 Hz. Here the positioning error *d* is 1.5 µm with a standard deviation of 0.5 µm. Increasing the frequency leads to an increase in the positioning error.

It is noticeable that the position error e_y is significantly higher than e_x. This indicates higher friction in the *y*-direction. The effect is likely caused by deviations in the angle of the setup. On the one hand, even with a precise experimental setup, there are still small errors in the levelling, resulting in a gravitational force acting against the alignment. On the other hand, uneven distribution of adhesive between the wafers leads to tilting between the glass surfaces that affects the homogenous distribution of capillary forces. This effect has not yet been observed in the research carried out on micro components. Due to the smaller geometric dimensions, the error influence would be small relative to the force equilibrium.

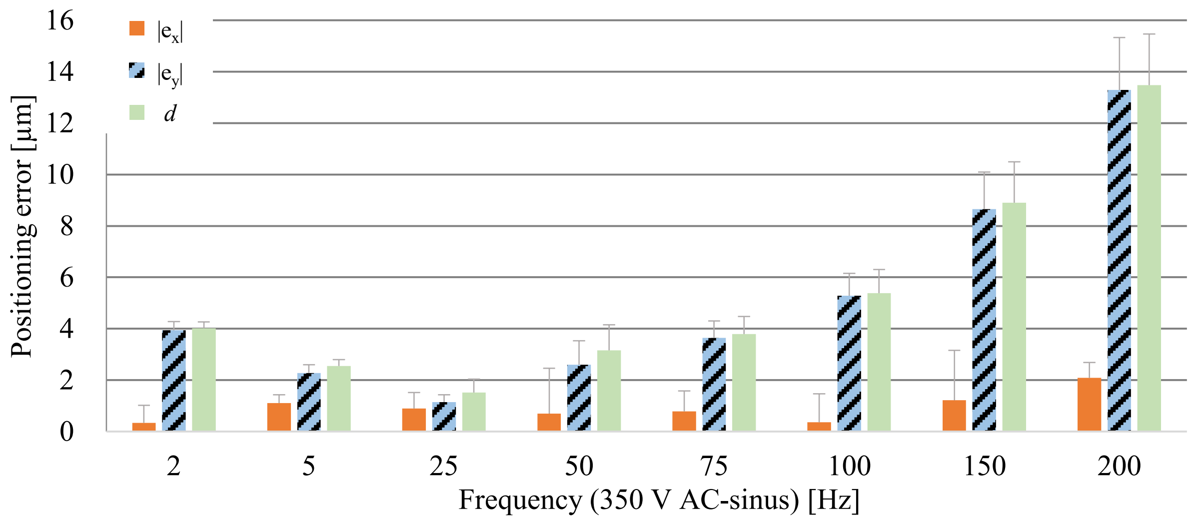


Figure 8: Dependence of the positioning error on the AC-frequency

6. Conclusion and outlook

Aligning components sequentially with high accuracy is time-consuming, expensive and requires special machine technology. In the past, there have been several approaches to address these problems. One of them is electrostatic self-assembly. We adopted this concept to develop assembly strategies for optical devices. This paper introduces a novel concept for full-wafer alignment with electrostatic self-assembly. The predicted application is the high-precision parallel assembly of optical devices.

The focus of this paper was to demonstrate the suitability of electrostatic self-assembly for large parts. Therefore, we combined the individual steps for a structure design in a consistent methodology. Using this methodology, we developed a structure design for the full-wafer alignment. In a new constructed experimental setup, we did several tests for the validation of the concept. During the test, we positioned a 125 mm glass wafer with a success rate of 100%. The experimental results prove that the electrostatic self-assembly, which was previously only used on millimetre-sized MEMS, can be successfully transferred to higher scales.

Two larger test series were carried out for this paper. In the first series, we achieved a positioning error *d* of 0.4 µm with 750 VAC, using a 105 mm alignment structure and 50 Hz alternating current. In the second series, we achieved a positioning error *d* of 1.5 µm with 25 Hz and 350 VAC. One additional finding is that on large scales, small deviations in the angle between the components have a significant influence on the positioning error and can lead to an uneven distribution of forces. To gather more data about the process, the distance between the glass surfaces must be measured in situ. Further development of the structure design also offers excellent potential. Specially designed pad arrangements can lead to better force distribution in a specific direction or block movement in specific directions. We plan to enhance the structure design with a corresponding simulation.

Acknowledgements

The research in this paper is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the Cluster of Excellence PhoenixD (EXC 2122, Project ID 390833453). The authors thank the German Research Foundation (DFG) for their support.

References

[1] Vinciguerra, S., Vinciguerra, M., 2019. Smart devices and healthy aging. NHA 5 (1), pp. 13–19.

[2] Blumenröther, E., Melchert, O., Kanngießer, J., Wollweber, M., Roth, B., 2019. Single Transparent Piezoelectric Detector for Optoacoustic Sensing-Design and Signal Processing. Sensors (Basel, Switzerland) 19 (9).

[3] Fricke, D., Denker, E., Heratizadeh, A., Werfel, T., Wollweber, M., Roth, B., 2019. Non-Contact Dermatoscope with Ultra-Bright Light Source and Liquid Lens-Based Autofocus Function. Applied Sciences 9 (11), 2177.

[4] Prossotowicz, M., Heimes, A., Flamm, D., Jansen, F., Otto, H.-J., Budnicki, A., Killi, A., Morgner, U., 2020. Coherent beam combining with micro-lens arrays. Optics letters 45 (24), pp. 6728–6731.

[5] Srinivasan, U., Helmbrecht, M.A., Muller, R.S., Howe, R.T., 2002. MEMS: Some Self-Assembly Required. Optics & photonics news 13 (11), pp. 20–25.

[6] Tondorf, M., Wilde, J., 2015. Elektrostatische und fluidische Self-Assembly-Prozesse für die präzise Montage von Mikrosystemen: Schlussbericht zu dem IGF-Vorhaben. Technische Informationsbibliothek u. Universitätsbibliothek, Freiburg. doi:10.2314/GBV:832344648.

[7] Dalin, J., Wilde, J., Lazarou, P., Aspragathos, N., 2011. Self-assembly of dies through electrostatic attraction: modelling of alignment forces and kinematics. J. Micro-Nano Mech. 6 (1-2), pp. 23–31.

[8] Dalin, J., Wilde, J., Zulfiqar, A., Lazarou, P., Synodinos, A., Aspragathos, N., 2010. Electrostatic attraction and surface-tension-driven forces for accurate self-assembly of microparts. Microelectronic Engineering 87 (2), pp. 159–162.

[9] Horechyy, A., Nandan, B., Zafeiropoulos, N.E., Formanek, P., Oertel, U., Bigall, N.C., Eychmüller, A., Stamm, M., 2013. A Step-Wise Approach for Dual Nanoparticle Patterning via Block Copolymer Self-Assembly. Adv. Funct. Mater. 23 (4), pp. 483–490.

[10] Whitesides, G.M., Grzybowski, B., 2002. Self-assembly at all scales. Science (New York, N.Y.) 295 (5564), pp. 2418–2421.

[11] Tondorf, M., Gan, Y., Mouselimis, K., Wilde, J., 2014. Elektrostatisch-fluidische Selbstassemblierung für die hochgenaue Mikro-Montage von MEMS. Elektronische Baugruppen und Leiterplatten DVS Berichte, pp. 46–51.

[12] Bohringer, K.-F., Goldberg, K., Cohn, M., Howe, R., Pisano, A., 1998. Parallel microassembly with electrostatic force fields, in: 1998 IEEE International Conference on Robotics and Automation. IEEE International Conference on Robotics and Automation, Leuven, Belgium. 16-20 May 1998. IEEE, Piscataway, pp. 1204–1211.

[13] Cohn, M.B., Boehringer, K.F., Noworolski, J.M., Singh, A., Keller, C.G., Goldberg, K.A., Howe, R.T., 1998. Microassembly technologies for MEMS, in: Micromachining and Microfabrication Process Technology IV. Micromachining and Microfabrication, Santa Clara, CA. Sunday 20 September 1998. SPIE, pp. 2–16.

[14] Dalin, J., Wilde, J., 2009. Self-assembly of micro-parts using electrostatic forces and surface tension, in: 59th Electronic Components and Technology Conference, 2009. ECTC 2009 ; San Diego, CA, USA. 5/26/2009 - 5/29/2009. IEEE, Piscataway, NJ, pp. 1517–1524.

Biography

Martin Stucki (*1986) studied mechanical engineering at Leibniz University Hannover (LUH) and graduated with a diploma in 2014. Since 2015, he has been a research associate at the Institute of Assembly Technology at LUH (match) where he researches in the field of precision assembly.

Christoph Schumann (*1990) studied mechanical engineering at Leibniz University Hannover (LUH) and graduated with a master's degree in 2018. Until 2021, he was a research associate at the Institute of Assembly Technology at LUH (match) where he researched in precision assembly and automated handling.

Annika Raatz (*1971) has been the head of the Institute of Assembly Technology (match) since its establishment in 2013. Prof. Dr.-Ing. Annika Raatz is a member of the Society for Assembly, Handling and Industrial Robotics (MHI) and the Academic Association for Production Technology (WGP). Her main research topics are robot aided handling processes, machine concepts for handling, assembly and production automation and soft material robotic systems.

2nd Conference on Production Systems and Logistics

Development of a Methodology for the Digital Representation of Manufacturing Technology Capabilities

Tobias Köhler[1,2], Ludwig Kleinhenz[1], Philipp Matthias Schäfer[1], Jean Pierre Bergmann[2], Diana Peters[1]

[1]*German Aerospace Center (DLR), Institute of Data Science, Jena, Germany*
[2]*Technische Universität Ilmenau, Production Technology Group, Ilmenau, Germany*

Abstract

The demand for efficient and digital systems for supporting the decision making during the design of a product is a key issue in manufacturing companies. Decisions made during the development and design of a product have a strong impact on the costs and delivery times of a product. Hence, a digital system which supports the engineer during and after the development process with information about the manufacturability of the product can reduce the production costs and times. In order to be able to evaluate the manufacturing capabilities at manufacturing process and machine level, there is a need to represent them in a digital way. Digital knowledge bases like taxonomies and ontologies provide the possibility for a representation of manufacturing resources. The state of the art shows different approaches for the use of ontologies in the domain of subtractive manufacturing processes as well as additive manufacturing (AM) processes. The goal of this work is the semantical representation of manufacturing technology capabilities with focus on AM-machines and processes. In this paper we introduce taxonomies of Manufacturing Features and Manufacturing Restrictions which were developed in accordance with current standards. To enrich the taxonomies with information, it was enhanced by relations between different manufacturing related entities in a knowledge graph. If manufacturing processes and machines can be digitally mapped, described and linked to the geometric information of a product together with information on the current performance of the company/network, bottlenecks and delivery delays during the manufacturing of parts can be avoided.

Keywords

Knowledge Representation; Taxonomy; Knowledge Graph; Ontology; Additive Manufacturing

1. Introduction

The ongoing demand on companies to reduce manufacturing costs and delivery times of products requires the development of new approaches and methods [1], [2], [3]. Decisions made during the design process of products have significant impact on costs as well as production and delivery times [4]. Incomplete information about the availability and capabilities of manufacturing technology resources in the factory can lead to design changes which can then cause delayed delivery times and increased cost [5], [6]. The use of information on available resources and capabilities in the company during the design phase of a product can prevent these problems [7]. Especially for the development of components with batch size one (engineer-to-order products), the development costs play a significant role. For small batch sizes innovative manufacturing processes such as additive manufacturing (AM) play an increasingly important role. In order to be able to evaluate the manufacturing possibilities of a factory automatically at the process and machine level, it is necessary to digitally represent the capabilities of manufacturing processes and manufacturing machines. Hence, these properties and capabilities must be described semantically and in a machine-readable manner. Furthermore, in addition to the requirements mentioned above, such a system must also be easy to extend, be used for the description of as many manufacturing processes as possible, and be used specifically for the description of component properties as well as machine capabilities. In this work, we represent the knowledge and information about the properties of a certain manufacturing process or the capabilities of a certain manufacturing machine digitally and semantically in an ontology.

There are various methods described in the literature for semantically representing knowledge [8], [9], [10], [11]. Those methods differ in quantity of richness of information and relationships. The semantic staircase model compares and sorts these models depending on their semantic richness [8]. In ascending order, the following models are on the semantic staircase: glossaries, taxonomies, thesauruses, topic maps, and ontologies. According to Blumauer and Pellegrini, ontologies are considered to be the richest. [8]

Gruber describes ontologies as follows: „An ontology is a formal, explicit specification of a shared conceptualization" [12]. This means that an ontology is used to describe the common understanding of reality. Furthermore, an ontology can be defined as the explicit formal definition of concepts in a particular setting and the relationships between them [13]. This formality allows knowledge to be represented in machine-readable form. [14], [15], [16]

Ontologies in manufacturing engineering exist for both the subtractive and additive manufacturing domains. The current state of the art shows, that most approaches are aiming at automated process step planning. For example, in the domain of subtractive manufacturing processes, Ma et al. [11] represent in their work the fundamental Manufacturing Features for the turning process in a taxonomy in order to be able to use them subsequently for manufacturing planning. In the area of additive manufacturing, Sanfilippo et al. [17] present an ontology that contains the fundamental principles of additive manufacturing. In addition to domain-specific ontologies, cross-domain knowledge representations for manufacturing processes are also presented by some authors (e.g. MASON [18] or MDSL [19]). Summarized the state of the art shows some gaps in the area of automated feasibility checks depending on capabilities of available additive manufacturing machines. The goal of this work is therefore to semantically describe additive manufacturing processes and the capabilities of manufacturing machines in an ontology in order to subsequently make a statement about the manufacturability of a component. A digital and semantic representation of manufacturing technologies and conditions in combination with evaluation algorithms, can be used to automatically evaluate a designed 3D model according to specifications of manufacturability and cost requirements. This approach is based on the idea that specific statements about the manufacturability on a certain machine can be made, by the combination of taxonomies for Manufacturing Features (MF) and Manufacturing Restrictions (MR) in an ontology. In future works, we will present a reasoner, which access the information represented in the ontology.

2. Literature review

The state of the art shows different existing approaches for this purpose, which will be reviewed in detail in the following sections. Current research can be divided into the broad domains of subtractive and additive manufacturing technologies. Furthermore, different goals can be pursued with the presented ontologies, for example, they can be used for process step planning, for the analysis of manufacturability or for the assignment of components with the corresponding manufacturing processes and/ or machines.

The following publications can be attributed to the field of process step planning of subtractive manufacturing processes. In 2015, Rehage and Gausemeier [20] presented the ontology *InVor*. In this ontology, tools of CNC machines are represented and the possible NC functions are assigned to them. The authors also follow the approach of determining alternative CNC machines based on the NC codes of a component mapped in the ontology. A similar approach for process step planning for milled parts is pursued by the authors Eum et al. [21]. The content of their publication is a system based on ontologies, in order to determine and assign the process steps of the manufacturing of a component. For this purpose, selected features for the milling process and their description by logical rules are explained in their paper. A publication on process planning of prismatic parts using a process model based on ontologies is presented by the authors Kang et al. [22] In this work, part features and manufacturing process steps which can be performed with a milling machine are described and then linked by appropriate rules.

In order to represent the properties of manufacturing technologies in an ontology, a taxonomy of possible MFs can be created, which contains geometries that can be manufactured by using individual technologies. This approach is pursued by Ma et al. [11] in the field of rotation-symmetric components. In their work, they represent the basic MFs for the turning process in a taxonomy in order to subsequently enable the planning of manufacturing. Besides subtractive manufacturing processes, ontologies for process step planning of additive manufacturing are also part of recent scientific research. For example, Liang [23] presents in his paper a methodology for representing additive manufacturing using a knowledge base. The result of the presented research is an ontology (*AM-OntoProc*) which aims at supporting process planning of additive manufacturing. Another part of current research is the automated assignment of components to processes and machines that can produce this component. Publications exist here in the area of both subtractive and additive manufacturing processes. Mesmer and Olewnik [24] published in their work a methodology based on an ontology model for the assignment of components and the corresponding subtractive manufacturing processes. The concept of the methodology relies on the fact that a manufacturing process can be determined based on features and attributes of a product. The objective of the publication by Ming et al. [25] is to automate and thus facilitate the selection of tools in CNC milling machines. In their paper, a taxonomy is presented, which is extended by relations and properties. The authors Sarkara and Šormaz [26] present in their work an ontology that describes the properties of manufacturing processes and matches them with the capabilities/boundaries of manufacturing resources. In their work, subtractive manufacturing processes are considered and represented. Part features are assigned to the corresponding subtractive manufacturing processes, and then the part features and manufacturing process steps are linked to the corresponding machines using logical rules to enable step-by-step process planning. In 2015, the authors Eddy et al. [27] presented an ontology that works in combination with a rule base. The use of the rule base allows, on the one hand, support for decision making in process planning in additive manufacturing and, on the other hand, determination of the appropriate manufacturing processes for a component. Furthermore, Kim, Rosen et al. [28] present the structure of an ontology in their 2019 publication and show how rules for creating AM components can be digitally described. In this ontology, MFs are introduced, which allow the representation of design features and their restrictions caused by the respective AM process and thereby achieve a manufacturability check with the help of these features. Ko et al. [29] present the extension of this work in 2021. The authors present a knowledge graph that supports engineers in the design of AM components.

The described studies show that there is a huge interest in the topic of knowledge representation for manufacturing technologies. For the more specific approach of assigning capable manufacturing machines, some authors describe the capability of machines and processes by means of MFs and restriction properties of a part. For those studies, some research gaps can be outlined. First, a manufacturability analysis, as well as an analysis of machine resolution and accuracy restrictions of additive manufacturing parts in order to assign a part to a certain machine individual is not considered. Second, the taxonomies of MFs were not compiled according to current standards of manufacturing technology, in order to be able to guarantee a reliable assignment of the features to the machines. Third, the Manufacturing Restrictions, which describe the capabilities of a machine and define the properties of the MFs were also not selected based on current standards for additive manufacturing. Furthermore, those restrictions were not introduced in a taxonomy model to provide a framework for ensuring an easy expandability.

3. Methodology

For the semantic description of knowledge of the capabilities of manufacturing processes and machines, a knowledge base is developed and presented in this paper. The capabilities of manufacturing processes are described by a semantic representation of manufacturable part geometries. These part geometries are further referred to as Manufacturing Features. In the presented method, the MFs are defined based on current manufacturing technology standards and then semantically represented in a taxonomy. The challenge here is to model the MFs in a way they represent a wide range of manufacturing processes. To describe the capabilities of a manufacturing process, the MFs are assigned to them in an ontology. For example, a turning process can only produce rotation-symmetric parts/features, whereas a milling process or an additive manufacturing process can also produce asymmetric features. Thus, a semantic description of manufacturable geometries of a process is provided. However, in order to be able to represent the capabilities of manufacturing machines in the same manner, it is not sufficient to assign only MFs. In reality, different machine individuals are able to execute the same manufacturing processes, but are not necessarily able to produce the same individuals of particular MFs. It is therefore necessary to define a set of restrictions that characterize both the capabilities of a machine and the geometric details of the MFs. These restrictions are called Manufacturing Restrictions (MR) in the following and are also defined according to current standards of manufacturing technology and represented in a taxonomy. In the section below, the structure of the ontology as well as the MFs and MRs are further described.

3.1 Knowledge Base

An ontology is used as the preferred knowledge base in this work, since this form of knowledge representation contains the most information. Furthermore, it is easily possible to extend the knowledge base e.g. with further MFs and MRs. The structure of the developed ontology is divided into two parts. The concepts of the manufacturing domain are represented in the Terminology Box (TBox). Here, the blueprints of the five main classes *ManufacturingFeature*, *ManufacturingRestriction*, *ManufacturingProcess, ManufacturingMachine* and *Part* are described with their corresponding subclasses and properties. The individuals of the classes are mapped in the Assertion Box (ABox). In addition to the part properties defined by the MFs and -Restrictions, the knowledge about the capabilities and limitations of particular machine individuals is also described here.

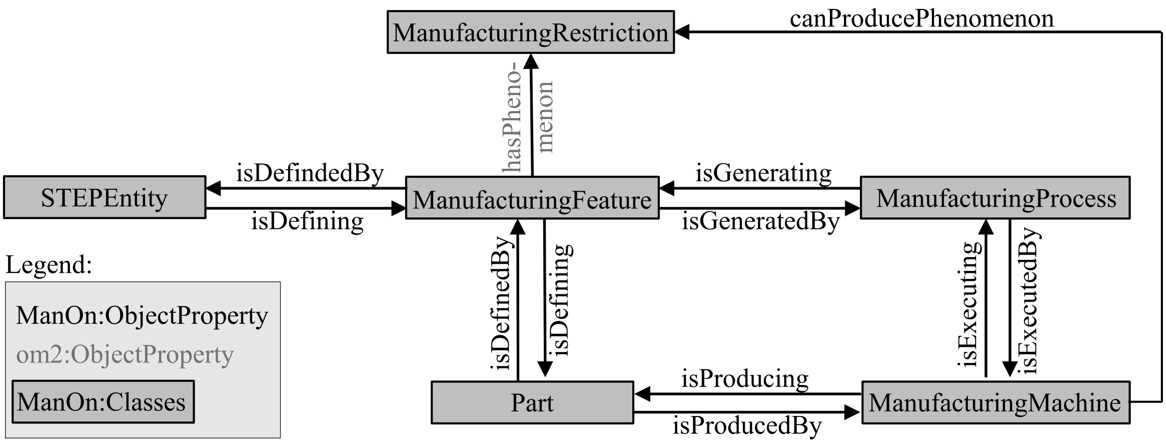


Figure 1: Structure of the presented ontology for providing manufacturability analyses

Figure 1 shows the TBox of the ontology. The shown structure of the ontology contains elements from the Ontology of units of Measure (om2) [30] in addition to self-defined classes and properties. At the center of the presented ontology is the class *ManufacturingFeature*, this class defines the class Part and is defined by class *STEPEntity*. *STEPEntity* describes the entities of the STEP-schema [31], which provides a standard for the exchange of product manufacturing data. Furthermore, the class *ManufacturingProcess* is described in the ontology. This class can generate a *ManufacturingFeature* and is executed by the class *ManufacturingMachine*. In order to be able to make a statement about the manufacturability of a *Part* in terms of accuracy requirements, geometric requirements and surface quality requirements, the *ManufacturingRestriction* class is linked to both *ManufacturingFeature* (by the ObjectProperty *om2:hasPhenomenon*) and *ManufacturingMachine* (by the ObjectProperty *canProducePhenomenon*). In the ontology it is defined which *ManufacturingFeature* has which *ManufacturingRestriction*. The properties associated with the *ManufacturingRestrictions* of features and machines can be compared, by the use of a reasoner, to achieve an assignment of MFs and Machines. The manufacturing technology aspect behind the *ManufacturingFeature* and *ManufacturingRestriction* classes is described in the next two sections.

3.2 Manufacturing Features

Before taking the step of semantically representing various manufacturing processes, the capabilities of those processes must first be analyzed and defined in more detail. A wide range of standards [32], [33], [34], [35] were consulted for this purpose in order to achieve a unified model of manufacturing process representation. Moreover, the Manufacturing Features, defined according to current manufacturing technology standards, are then used to build up a taxonomy (see Figure 2). The represented information about manufacturable geometries can be used for describing the manufacturing process capabilities.

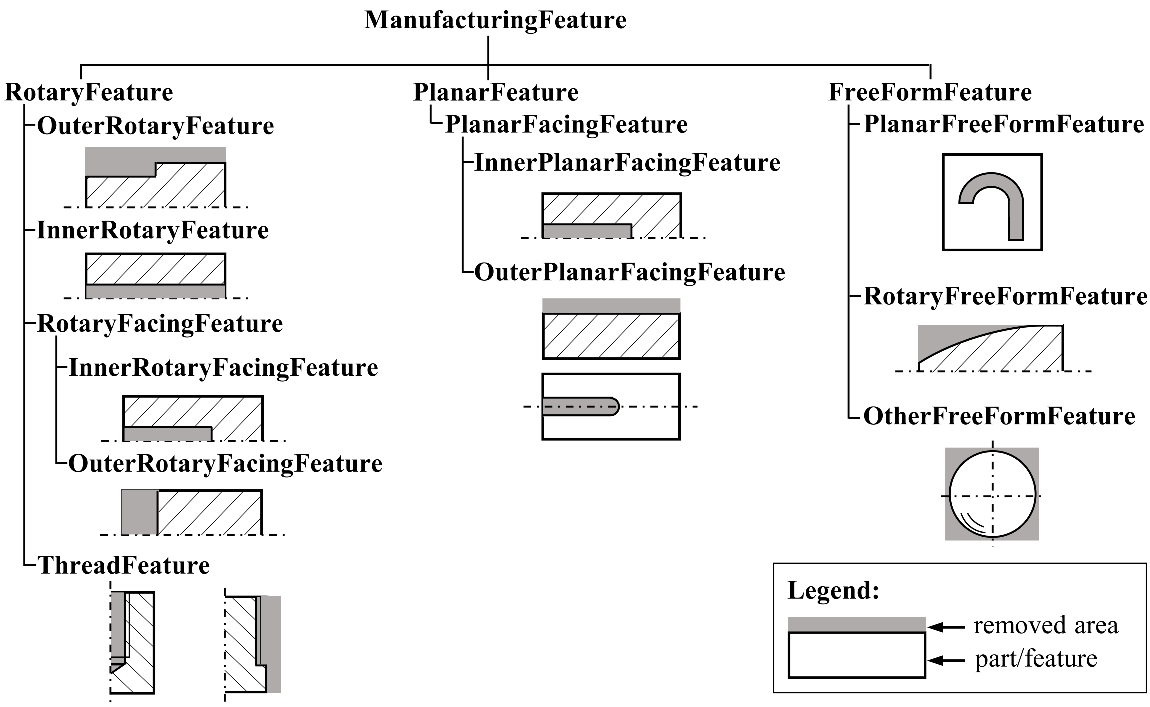


Figure 2: Representation of the taxonomy for Manufacturing Features with selected illustrated geometries

Based on the standard DIN 8589 [36], the MFs are defined in the following sections and specified in a taxonomy (see Figure 2). Since it is possible to produce almost all geometric shapes with additive manufacturing processes, subtractive manufacturing standards were used for modeling the MF taxonomy in order to obtain a more detailed description. The taxonomy of MFs can be used to define both, subtractive and additive manufacturing processes, by combining the MFs with the MRs that have been modelled on the basis of current additive manufacturing standards (see chapter 3.3).

In the presented taxonomy, a MF is divided into the three areas of geometric shapes: rotation-symmetric features (*RotaryFeature*), planar features (*PlanarFeature*) and other features (*FreeFormFeature*). The *FreeFormFeatures* here represent both rotation-symmetric features and planar features with a freely designed shape. This classification relies on the capabilities of different manufacturing processes. For example, all subclasses of *RotaryFeature* can be produced by a turning process.
These features are further divided by a classification based on the location of the machining area. The subclasses of *RotaryFeature* are divided into inner (*InnerRotaryFeature*) and outer machining (*OuterRotaryFeature*), as well as thread machining (*ThreadFeature*) and face machining (inner and outer machining) (*RotaryFacingFeature*). *PlanarFeature* is also divided into inner and outer machining. An *InnerPlanarFacingFeature* can be defined by the base area of a rotation-symmetric hole, a pocket or a slot. *OuterPlanarFacingFeature* refers to all geometric elements that are located on the outer surface of a component. This feature can represent geometries such as walls and other planar contours. Since the *FreeFormFeature* class describes a mixture of planar free form contours (*PlanarFreeFormFeature*), rotation-symmetric free form contours (*RotaryFreeFormFeature*) and other free form contours (*OtherFreeFormFeature*), all geometric elements that embody complex geometries are covered in this class.

3.3 Manufacturing Restrictions

The layered build-up of parts by additive manufacturing processes makes it possible to produce almost any geometric shape. Hence, a semantic process description only using MFs is not possible. In addition to the geometric description of the MFs, the limitations of manufacturing machines must also be described semantically and represented in a taxonomy. Those limitations are called Manufacturing Restrictions in this work. For example, it depends on the process type whether and to what extent unsupported or overhanging structures can be manufactured. Furthermore, restrictions in the area of tolerances and surface qualities have to be included.
Figure 3 shows the taxonomy of MRs. This taxonomy was developed by including standardized test characteristics for restricting geometrical properties as well as standardized technical aspects for the selection of suitable additive manufacturing processes. These were developed based on the guideline VDI 3405 sheet Nr. 3.2 [35], the standard DIN EN ISO/ASTM 52902 [33] and the standard DIN EN ISO/ASTM 52910 [34]. Furthermore, the elements of geometric tolerancing according to DIN EN ISO 1101 [32] were included in the taxonomy and described semantically. MRs can be different, with the same process, depending on the type of machine. In the presented taxonomy, the class of MRs is divided into the subclasses *Accuracy*, *Resolution*, *Manufacturability* and *SurfaceStructure*. The subclasses of these classes are described in Figure 3 with example drawings. In the following section, these subclasses will be further explained.
The accuracy with which a machine can produce a MF can be a decisive limitation for the selection of the machine. These limitations are represented in the *Accuracy* class, which is divided into *SizeTolerance*, *GeometricTolerancing* and *GeometricDimensioning*. The last two subclasses are described in accordance with DIN EN ISO 1101 [32]. According to this standard, *GeometricTolerancing* describes, for example, the roundness of a cylinder or the flatness of a surface. Examples of *GeometricDimensioning* are the parallelism between two planes or the concentricity of two cylinders. Another point in the area of restriction *Accuracy* is *SizeTolerance*. This restriction describes how precisely a machine can reproduce a digital model or drawing, or how strong the deviations from the model or drawing are.
The minimal manufacturable geometric structures are described by the MR of the maximum *Resolution* of a machine. Depending on the process and machine type, there is a minimum size of a feature, which can be created by the size of the tool or by the minimum layer width. Walls defined by a minimum *WallThickness* and pins defined by a minimum cylindrical *PinDiameter* can be assigned to the positive volume range of a part. In turn, the restrictions minimum *HoleDiameter* and minimum *GapWidth* are assigned to the area of

negative volume. The minimum *EdgeRadius* describes the minimum manufacturable radius of an inner or outer edge.

The *Manufacturability* restrictions presented are mainly related to additive manufacturing, since this paper focuses on the semantic representation of AM-technologies. Some additive manufacturing processes make it possible to create cantilevered structures and overhangs by building removeable support structures. The restriction of the maximum overhang is described in the taxonomy by the class *UnsupportedStructure*. The minimum *TiltAngle* describes the minimum angle at which a structure can be situated in relation to the base plane. Above a certain tilt angle, a support structure is necessary in additive manufacturing to be able to produce a MF. Furthermore, the *HorizontalHoleDiameter* class describes the minimum producible diameter of a hole that is horizontal to the base surface.

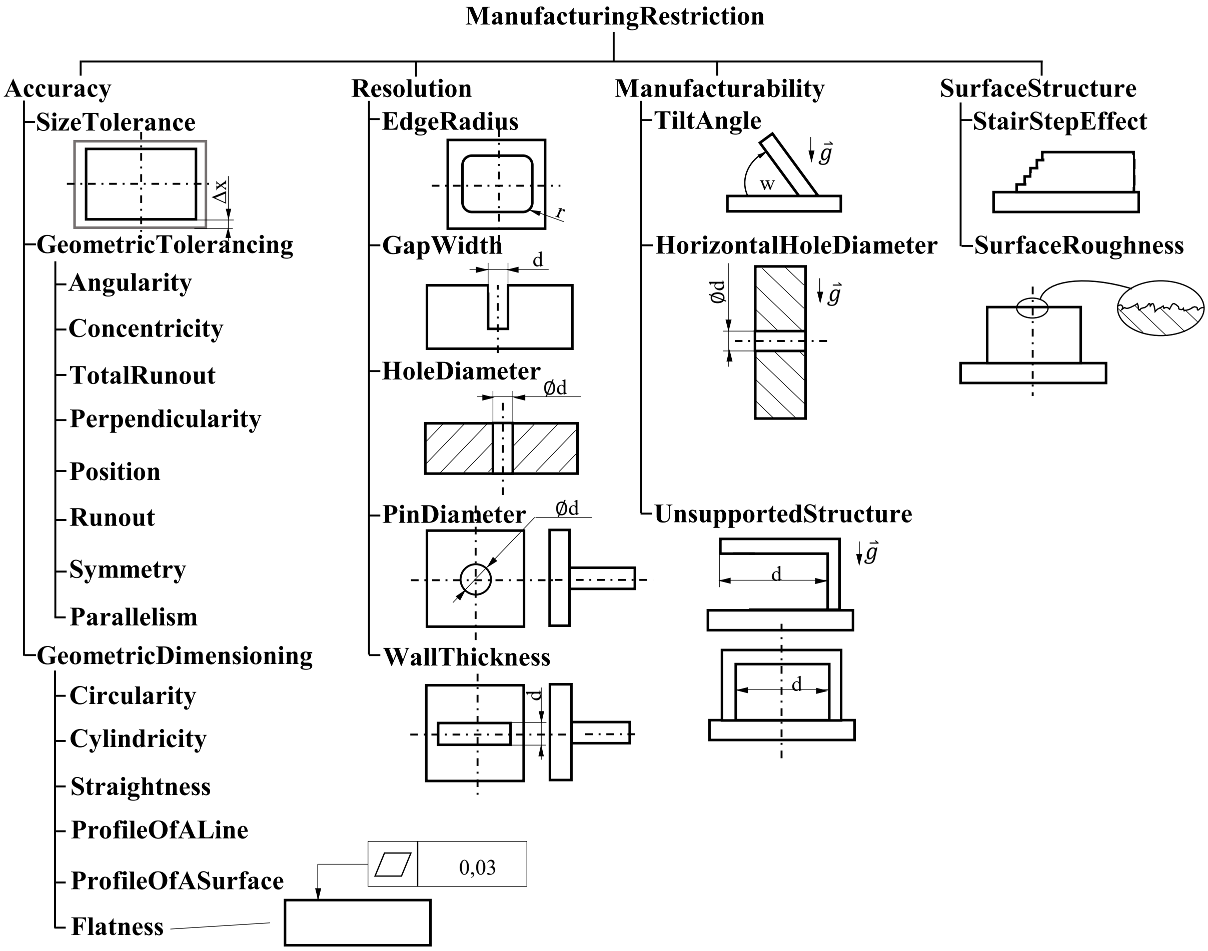


Figure 3: Representation of the taxonomy for Manufacturing Restrictions with selected illustrated geometries

Depending on the manufacturing process and the machine used, different qualities of surface structure can be generated. Those restrictions are represented in the class *SurfaceStructure*. The *SurfaceRoughness* applies as a restriction for both subtractive and additive manufactured components. If a function is assigned to a certain surface, then a maximum tolerated surface roughness must be defined. However, if a machine cannot achieve this quality, this has a limiting effect on component production. In addition to *SurfaceRoughness*, there is a further restriction in the area of surface structure. Due to the layered structure of a component in additive manufacturing, a step effect occurs horizontally to the build direction. This effect depends on the minimum layer height of a manufacturing process and the tilt angle of a feature. For example, a powder-bed laser process has a lower minimal layer height than a wire-based process. In the taxonomy this restriction is represented in the class *StairStepEffect*.

4. Conclusion

In this work we present a method to digitally and semantically represent geometric properties of components as well as the capabilities of additive manufacturing processes and machines. The basis of this method is an ontology that provides the structure for matching the required properties of components with the existing capabilities of machines. The representation of parts is done by geometric MFs in combination with limiting MRs. The described MFs combined with the MRs completely define and represent a part on an abstract level to achieve the goal of semantically modeling manufacturing processes. MRs are also used to describe the capabilities of manufacturing machines. The taxonomies presented in this paper contain the geometrical description of components (MFs) as well as machine and geometry dependent restrictions (Manufacturing Restrictions). The MF taxonomy is categorized according to the criteria of geometric shape (rotation-symmetric, planar and other contoured). This taxonomy can be applied to represent additive as well as subtractive manufacturing processes. With the MR taxonomy, we present a taxonomy for the limiting restrictions for additive manufacturing. In future works, we will extend and apply the MR taxonomy to subtractive manufacturing. In addition to the expansion of new process types and machine types, changes due to innovations in the field of manufacturing technology result in new Manufacturing Features and -Restrictions. Due to the structure of the taxonomies and the ontology, new restrictions and features can easily be added and extended. In the coming future the results of this paper are used to build a knowledge graph, which is used to plan the capacity utilization of a factory automatically, by assigning parts to a capable machine. Those automated assignments result from a comparison of the required MRs of a part and the existing restrictions of the machines in a factory. Furthermore, we will validate and evaluate the presented methodology with real world data.

References

[1] Cohen, Y., Faccio, M., Pilati, F., Yao, X., 2019. Design and management of digital manufacturing and assembly systems in the Industry 4.0 era. Int J Adv Manuf Technol 105 (9), 3565–3577.

[2] Reimche, M., Berghof, S., Bergmann, J.P., 2021. Adapted Process Model for Manufacturing Within Production Networks, in: , Production at the leading edge of technology, vol. 49. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 611–620.

[3] Schuh, G., Prote, J.-P., Gützlaff, A., Thomas, K., Sauermann, F., Rodemann, N., 2019. Internet of Production: Rethinking production management, in: Wulfsberg, J.P., Hintze, W., Behrens, B.-A. (Eds.), Production at the leading edge of technology. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 533–542.

[4] Feldhusen, J., Grote, K.-H., 2013. Pahl/Beitz Konstruktionslehre. Springer Berlin Heidelberg, Berlin, Heidelberg, 883 pp.

[5] Kuhlmann, T., 1994. Konzeption und Entwicklung eines Systems zur Koordinierung der Produktion komplexer Unikate. Hochschulschrift, Bremen.

[6] Eckstein, H., Eichert, J., Waidmann, J., 2010. Prozessmodell zur Integration von Konstruktion und Montageplanung. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb 105 (3), 200–205.

[7] Schützer, K., Bertazzi, J.d.A., Sallati, C., Anderl, R., Zancul, E., 2019. Contribution to the development of a Digital Twin based on product lifecycle to support the manufacturing process. Procedia CIRP 84, 82–87.

[8] Blumauer, A., Pellegrini, T., 2006. Semantic Web und semantische Technologien: Zentrale Begriffe und Unterscheidungen, in: Pellegrini, T., Blumauer, A. (Eds.), Semantic Web. Wege zur vernetzten Wissensgesellschaft. Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, pp. 9–25.

[9] Cao, Q., Zanni-Merk, C., Reich, C., 2019. Ontologies for Manufacturing Process Modeling: A Survey, in: Dao, D., Howlett, R.J., Setchi, R., Vlacic, L. (Eds.), Sustainable Design and Manufacturing 2018, vol. 130. Springer International Publishing, Cham, pp. 61–70.

[10] Keil, J.M., Schindler, S., 2018. Comparison and evaluation of ontologies for units of measurement. SW 10 (1), 33–51.

[11] Ma, H., Zhou, X., Liu, W., Niu, Q., Kong, C., 2020. A customizable process planning approach for rotational parts based on multi-level machining features and ontology. Int J Adv Manuf Technol 108 (3), 647–669.

[12] Gruber, T.R., 1995. Toward principles for the design of ontologies used for knowledge sharing? International Journal of Human-Computer Studies 43 (5-6), 907–928.

[13] Gruber, T.R., 1993. A translation approach to portable ontology specifications. Knowledge Acquisition 5 (2), 199–220.

[14] Studer, R., Benjamins, V., Fensel, D., 1998. Knowledge engineering: Principles and methods. Data & Knowledge Engineering 25 (1-2), 161–197.

[15] Gilchrist, A., 2003. Thesauri, taxonomies and ontologies – an etymological note. Journal of Documentation 59 (1), 7–18.

[16] Uschold, M., Gruninger, M., 1996. Ontologies: principles, methods and applications. The Knowledge Engineering Review 11 (2), 93–136.

[17] Sanfilippo, E.M., Belkadi, F., Bernard, A., 2019. Ontology-based knowledge representation for additive manufacturing. Computers in Industry 109, 182–194.

[18] Lemaignan, S., Siadat, A., Dantan, J.-Y., Semenenko, A., 2006. MASON: A Proposal For An Ontology Of Manufacturing Domain, in: IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06). IEEE Workshop on Distributed Intelligent Systems: Collective Intelligence and Its Applications (DIS'06), Prague. 2006. IEEE, pp. 195–200.

[19] Ameri, F., Dutta, D., 2006. An Upper Ontology for Manufacturing Service Description. Proceedings of IDETC/CIE.

[20] Rehage, G., Gausemeier, J., 2015. Ontology-based Determination of Alternative CNC Machines for a Flexible Resource Allocation. Procedia CIRP 31, 47–52.

[21] Eum, K., Kang, M., Kim, G., Park, M.W., Kim, J.K., 2013. Ontology-based modeling of process selection knowledge for machining feature. Int. J. Precis. Eng. Manuf. 14 (10), 1719–1726.

[22] Kang, M., Kim, G., Lee, T., Jung, C.H., Eum, K., Park, M.W., Kim, J.K., 2016. Selection and sequencing of machining processes for prismatic parts using process ontology model. Int. J. Precis. Eng. Manuf. 17 (3), 387–394.

[23] Liang, J.S., 2018. An ontology-oriented knowledge methodology for process planning in additive layer manufacturing. Robotics and Computer-Integrated Manufacturing 53, 28–44.

[24] Mesmer, L., Olewnik, A., 2018. Enabling supplier discovery through a part-focused manufacturing process ontology. International Journal of Computer Integrated Manufacturing 31 (1), 87–100.

[25] Ming, Z., Zeng, C., Wang, G., Hao, J., Yan, Y., 2020. Ontology-based module selection in the design of reconfigurable machine tools. J Intell Manuf 31 (2), 301–317.

[26] Sarkar, A., Šormaz, D., 2019. Ontology Model for Process Level Capabilities of Manufacturing Resources. Procedia Manufacturing 39, 1889–1898.

[27] Eddy, D., Krishnamurty, S., Grosse, I., Perham, M., Wileden, J., Ameri, F., 2015. Knowledge Management With an Intelligent Tool for Additive Manufacturing, in: Volume 1A: 35th Computers and Information in Engineering Conference. ASME 2015, Boston, Massachusetts, USA.

[28] Kim, S., Rosen, D.W., Witherell, P., Ko, H., 2019. A Design for Additive Manufacturing Ontology to Support Manufacturability Analysis. Journal of Computing and Information Science in Engineering 19 (4).

[29] 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.

[30] Rijgersberg, H., van Assem, M., Top, J., 2013. Ontology of units of measure and related concepts. SW 4 (1), 3–13.

[31] International Organization for Standardization, 2014. Industrial automation systems and integration: Product data representation and exchange 25.040.40.

[32] DIN EN ISO 1101:2017-09, Geometrische Produktspezifikation_(GPS)_- Geometrische Tolerierung_- Tolerierung von Form, Richtung, Ort und Lauf (ISO_1101:2017); Deutsche Fassung EN_ISO_1101:2017. Beuth Verlag GmbH, Berlin. doi:10.31030/2552752.

[33] DIN EN ISO/ASTM 52902:2020-05, Additive Fertigung_- Testkörper_- Allgemeine Leitlinie für die Bewertung der geometrischen Leistung additiver Fertigungssysteme (AM-Systeme); German Version EN_ISO/ASTM 52902:2019. Beuth Verlag GmbH, Berlin. doi:10.31030/3057421.

[34] DIN EN ISO/ASTM 52910:2020-02, Additive Fertigung_- Konstruktion_- Anforderungen, Richtlinien und Empfehlungen (ISO/ASTM 52910:2018); Deutsche Fassung EN_ISO/ASTM 52910:2019. Beuth Verlag GmbH, Berlin. doi:10.31030/3111254.

[35] Verein Deutscher Ingenieure e.V., 2019. Additive Fertigungsverfahren: Gestaltungsempfehlungen - Prüfkörper und Prüfmerkmale für limitierende Geometrieelemente. Beuth Verlag GmbH, Berlin ICS 25.030.

[36] DIN 8589-0:2003-09, Fertigungsverfahren Spanen_- Teil_0: Allgemeines; Einordnung, Unterteilung, Begriffe. Beuth Verlag GmbH, Berlin. doi:10.31030/9500664.

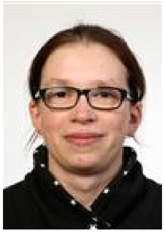
Biography

Tobias Köhler studied mechanical engineering at the Technische Universität Ilmenau, Germany. Since 2017, he has been a PhD candidate and research associate in the Production Technology Group (IFt) at the Technische Universität Ilmenau. Since 2019, he has been working for a scientific cooperation between IFt and the German Aerospace Center (DLR). His research topics are solid state welding processes and the digitization in production technologies.

Ludwig Kleinhenz studied mechanical engineering at the Technische Universität Ilmenau. Since 2020, he has been a research associate at the Institute of Data Science in the German Aerospace Center (DLR) with a research focus on digital production platforms and the digitization of production technologies.

Philipp Matthias Schäfer studied Mathematics at the Technical University of Ilmenau, the University of Cambridge, and at Ulm University, receiving his PhD in 2012. He worked for 6 years as a Software Developer. Since 2017 he is a research associate at the Institute for Data Science in the German Aerospace Center (DLR) working on ontologies for robotics.

Jean Pierre Bergmann studied mechanical engineering at the University of Ancona and received his Ph.D. (Dr.-Ing.) from the University of Bayreuth in 2003. He then worked as a PostDoc at the Technische Universität Ilmenau till 2007, moving then to JENOPTIK Automatisierungstechnik in Jena. He has been head of the Production Technology Group (IFt) at the Technische Universität Ilmenau since 2010. His research topics are joining technologies and the digitization of these.



Diana Peters studied computer science at the Friedrich-Schiller-University in Jena. From 2017 to 2018, she was a research associate at the Institute of Data Science in the German Aerospace Center (DLR). Since 2018, she has been team lead of the Digital Production Platforms Group at DLR. Her research interests are the digitization of manufacturing processes with focus at the space domain as well as modelling.

2nd Conference on Production Systems and Logistics

Prescriptive Maintenance for Onshore Wind Turbines

Bernhard Strack[1], Jana Frank[1], Volker Stich[1], Florian Pfau[1]
[1] *Institute for Industrial Management (FIR) at RWTH Aachen University, Aachen, Germany*

Abstract

Electricity generated by wind turbines (WT) is a pillar of the transition to renewable energy [1]. In order to economically utilize WTs, operating and maintenance costs, which account for 25% of total electricity generation costs in onshore WTs, are a focus of cost reduction activities [2]. A prescriptive maintenance approach can support in achieving this goal. Prescriptive maintenance is a maintenance approach, where asset condition data is collected and analyzed to recommend specific actions to prevent breakdowns and reduce downtimes. However, the processing and analysis of data is quite complex. Especially unstructured data (such as comments of service technicians in free text fields) is often left unused, as companies, mostly SMEs lack the capacity to carry out these analyses. In this work we propose an approach to utilize the information from service reports, maintenance reports as well as status records from SCADA systems for the development of a prescriptive maintenance approach to onshore WTs. To achieve this, an ontology was utilized in this approach to codify implicit knowledge of service technicians and aid in making unstructured data usable for further analysis. The ontology was used to link historical service and maintenance reports with status codes, thus enabling automated analysis. In interviews with WT topic experts and through further research, damage mechanisms and corresponding maintenance measures were identified and a measure catalogue was developed to support service and maintenance activities. The recognition of the root cause of problems allows for a prescriptive maintenance approach that recommends targeted actions to reduce downtimes and optimize maintenance activities, it also allows to effectively control the outcome of maintenance activities and optimize their execution.

Keywords

Prescriptive Maintenance; Ontology; Data Analytics

1. Introduction

Electricity generation from renewable energies continues to gain importance in Germany, surpassing electricity generation from fossil fuels for the first time in 2020 [3]. The largest share (41.7%) of electricity generation from renewable energies comes from onshore wind turbines (WTs) [4]. This highlights the key role of onshore WTs in achieving an environmentally friendly energy supply for Germany. One of the biggest challenges for operators of wind turbines (WTs) is the increasing cost pressure [5]. This is primarily caused by 2 factors: the expiration of the EEG (German: Erneuerbare Energien Gesetz; English: German renewable energy law) subsidies and increasing operating and maintenance costs (O&M costs) as the turbines age [6–8]. While the annual O&M costs of a new turbine are still at 10 to 15% of the average total costs per kWh, this share rises to as much as 35% by the end of its service life. On average, the O&M costs thus make up about 25% of the total costs per kWh [2]. This makes O&M costs a powerful lever for reducing the power generation costs of onshore WTs.

Maintenance activities include upkeep, inspection, repair and improvement [9] and usually take place at fixed intervals according to predefined maintenance specifications. This is where a promising new approach to reducing maintenance costs comes into play: the condition-based maintenance strategy [10]. Such a strategy is aimed at optimizing maintenance activities for time and personnel utilization, depending on the system and life cycle phase [10]. The basis of such maintenance optimization is detailed knowledge regarding the condition of the WT. For this purpose, large operators of wind farms or manufacturers of WTs evaluate operating and structured data from the maintenance history, but small and medium enterprises (SMEs) do not always have the capacities for such an evaluation. Regardless, comments on the condition of the turbine are often not used at all. These comments are often contained in non-standardized form in service reports and are documented by service technicians in free text fields. The latter are used to record the experience of service technicians in the otherwise standardized service reports. However, especially against the background of the increasing importance of interconnectedness and information processing in the yield and maintenance cost optimization of WTs, these sources of information are becoming more and more important. The possession of data and the ability to use it as a base for decisions on process optimization brings decisive cost advantages in the highly competitive energy sector.

This is exactly the lever, which the research project “ReStroK” targets. Together with partners from research and industry, the Institute for Industrial Management (FIR) at RWTH Aachen University is investigating the machine evaluation of both status data and maintenance histories, which have so far been used primarily for documentation purposes. The focus is on the avoidance of breakdowns due to progressive wear and tear of individual components by timely and effective maintenance of these components.

In the context of ReStroK a procedure for converting the information contained in the free texts of service reports regarding the turbine condition into a usable form was developed. This was realized, based on a previously developed specialized ontology [11]. The focus of this paper is however the damage mechanisms in onshore WTs and corresponding maintenance measures. In interviews with WT topic experts and through further research, failure modes were linked with optimized maintenance measures, and a measure catalogue was developed to support service and maintenance activities. In a first step, the analysis focused on the pitch system as well as the converter.

2. State of the art

2.1 Prescriptive maintenance

In the context of Industry 4.0, a distinction is made between different development stages. First, visualization and transparency are used to identify what is happening and why. Then prediction and finally self-optimization is made possible through data analysis and digital connectivity [12]. According to Davenport, within these stages of development, there are several levels of competitive advantages that can be achieved through business intelligence and analytics (Figure 1). The highest being the use of analytics for optimization of business activities, such as maintenance [13].

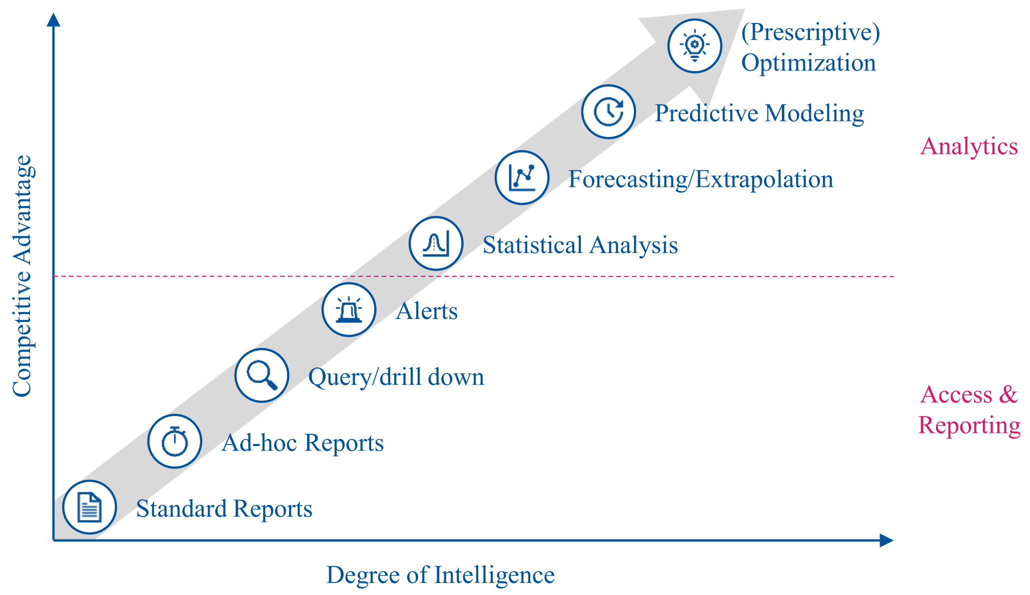


Figure 1: Business intelligence and analytics (Source: own representation based on Davenport [13])

Maintenance activities can be clustered in corrective (after fault occurrence) and preventive maintenance (prior to fault occurrence). Predictive maintenance approaches, a sub-set of preventive maintenance aim to carry out maintenance activities derived from data analytics and evaluation of significant system parameters [14]. However, predictive maintenance approaches are further differentiated. One such extension is the prescriptive maintenance approach. Whereas in a predictive maintenance approach only the condition of the object of maintenance is considered to carry out maintenance activities based on a forecast, a prescriptive maintenance approach considers the object of activities to be a guiding and controlling element for activities as well [15]. Such a maintenance approach would not just include the performance of maintenance activities prior to fault occurrence but also prescribe recommendations for specific actions to prevent breakdowns and reduce downtimes. A prescriptive maintenance approach, includes the collection of data of significant parameters, its analysis and evaluation and finally the utilization of these insights to prescribe activities to the object of the maintenance strategy. Based on data, measures are prescribed, evaluated for effectiveness on the object of maintenance activities and adapted when needed. This achieves the highest level of maturity and competitive advantages for a system, to be self-optimizing [13,16]. The basic data framework in such a prescriptive maintenance approach is shown in Figure 2.

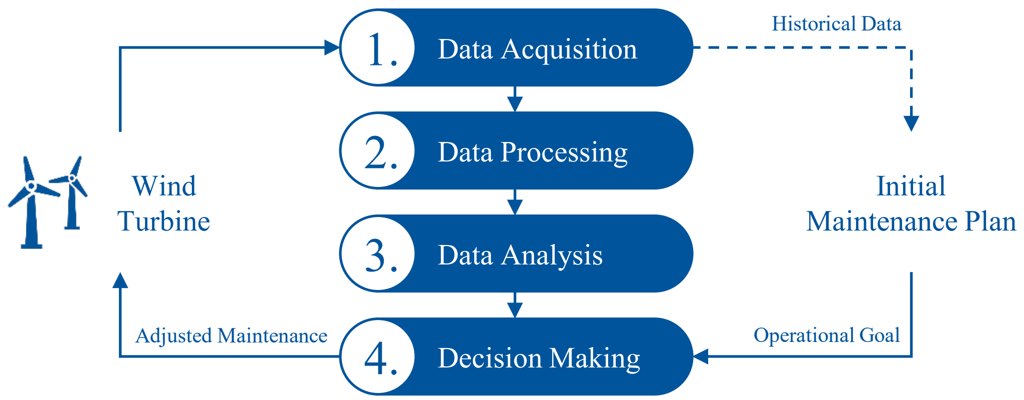


Figure 2: Data framework in a prescriptive maintenance approach (Source: own representation)

2.2 Utilization of data for maintenance of wind turbines

In order to implement a prescriptive maintenance approach with its inherent optimization characteristics, it is necessary to access and evaluate data of the object of activities [17]. Onshore WTs collect and transmit a variety of different data that can be used for prescriptive maintenance. These data are the master data of the WT, operating data sourced from Supervisory Control and Data Acquisition Systems (SCADA), which in part comprise hundreds of measuring channels, status and alarm messages as well as event and maintenance data, and data from the component-specific condition monitoring system (CMS) [18]. Previous approaches to optimize the maintenance of onshore WTs focus mainly on the analysis of data from the SCADA system and CMS. Comparatively little attention is paid to the maintenance data from service reports, which are at the center of the optimization approach developed in the context of "ReStroK".

The reasons for the nonuse of data are manifold. On the one hand, inadequate organizational structures may hinder the evaluation of already existing data, on the other hand, data may not be usable at all, since, for example, maintenance histories are often available in different formats such as paper or PDF. In addition, numerous different actors are involved in the various stages of a WT's lifecycle, further amplifying the problem of different data formats and data bases. In the context of prescriptive maintenance, the operating phase is of particular interest. This phase is characterized by a "service network" consisting of the operator, independent service providers (ISPs), OEMs and suppliers [19]. Especially OEMs and ISPs are often in direct competition with each other [20]. For this reason, there is usually little to no exchange of data and/or knowledge, which would be essential for targeted optimization of maintenance. As a result, hardly any technical system structures and standardized event classifications for the recording, storage and use of event data or information from service reports that are accepted throughout the industry exist. Conversations and interviews with OEMs and ISPs confirm, that software systems are mainly used for the documentation towards operators. Some specific concepts have already been developed and implemented together with expert committees, which, adapted to the needs of the wind energy industry, should enable cross-company work [10,21]. Examples of such standards are the internationally standardized reference designation system for power plants (RDS-PP), the state-event-root cause key for power generation units and the Global Service Protocol as a technical guideline of the Fördergesellschaft Windenergie (FGW). However, our research data shows, that from the OEMs and ISPs perspective the effort to implement the systems mentioned above is not always worthwhile. In addition, OEMs and ISPs often already have their own systems and the ability to perform in-house analysis. Accordingly, industry-wide cooperation and increased transparency is seen as a disadvantage for their own competitive position. This poses a major challenge for SMEs in particular, as they were often found not to have the appropriate capacities for in-house data analysis and maintenance optimization.

Faulstich and Hahn however, point out the added value of an evaluation of event data [18]. Event data on the one hand includes the automatically generated status messages of the WT and on the other hand maintenance data from service reports. Status codes report status changes of the WT and can thus indicate necessary maintenance measures. The corresponding status codes already provide information about the affected systems. Yet, often there is no clear assignment of a faulty condition to a causal component. Rather, as confirmed by interviews with ISPs, it is a "black box" containing a number of different components that may have caused the fault.

To overcome this challenge an ontology is needed. "An ontology is a formal, explicit specification of a shared conceptualization" [22]. In the specific use case for the service reports of onshore WTs, this means that the ontology is suitable as a link and common language between service technicians or maintenance reports and the SCADA system or status codes. Various approaches to ontologies can be found in the literature. For example, Zhou et al. developed an ontology based on Failure Mode and Effects and Criticality Analysis (FMECA) [23], Papadopoulos and Cipcigan developed an ontology that can be used for gearbox failure detection and diagnosis [24], and Ertek et al. performed an ontology-based analysis of the external circumstances and environmental conditions of WT accidents [25]. However, all these ontologies do not use service reports. Therefore, within the project "ReStroK" a new ontology was developed based on the evaluation of service reports that establishes a connection between the formal turbine structure, status data and free texts and the experience knowledge of the service technicians [11]. This ontology represents an important basis for the prescription of maintenance activities of onshore WTs and the necessary identification of damage mechanisms as well as corresponding maintenance measures.

3. Development of a prescriptive maintenance approach for onshore wind turbines

The access to asset data, including unstructured data from service reports is a prerequisite for the realization of a prescriptive maintenance approach [17]. However, especially SMEs often do not have access to such data. Thus, an ontology can be used to make such data accessible and usable [11]. Based on this ontology a prescriptive approach was developed as a part of the “ReStroK” research project to realize a prescriptive maintenance strategy. The main steps in the developed approach are the definition of a scope, the set-up of prescriptive maintenance steering and control, utilizing a previously developed ontology and the development of the function tree as well as the measure catalogue. The structure of the developed prescriptive maintenance approach is shown below in Figure 3.

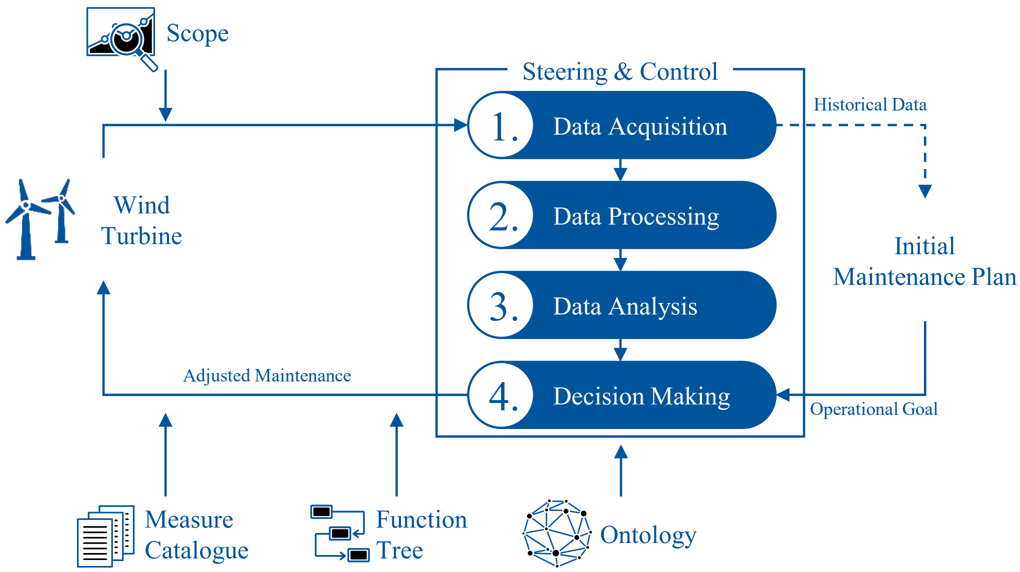


Figure 3: Prescriptive maintenance approach framework (Source: own representation)

3.1 Scope

While the overall functionality of onshore WTs does not differ greatly between manufacturers, the same cannot be said for their structure and subsystems. However after an analysis of the maintenance history of the WTs under consideration it was found that roughly 80% of the failures could be traced back to the pitch and to the converter system. Considering the time and capacity restrictions of the research project it was decided to put the initial scope for this work on the pitch and converter system.

3.2 Ontology to enable maintenance steering and control

As shown in Figure 3, data is needed to initiate the correct maintenance activities and control their effect. Concurrently it is used to make adjustments when needed. Especially unstructured knowledge from maintenance and service reports offers valuable insights into the state of WTs as well as the effect of conducted maintenance activities. Strack, Lenart et. al. developed an ontology, specifically tailored to onshore WTs, to link the implicit knowledge of service technicians with the operating and master data of

WTs and to enable the automated evaluation of service reports regarding fault findings and the evaluation of effects [11]. The developed ontology enables the incorporation of the unstructured knowledge found in service and maintenance reports for maintenance activities. This is made possible by linking error codes and maintenance reports. In the process of developing the ontology, the relevant terms were identified using text processing. This resulted in a database of terms and term frequencies, which were filtered in the next step. Finally, the terms were clustered and relations were defined to formalize the ontology. By using this ontology for this project, it is possible to automatically evaluate existing maintenance and service reports for issues and potential future breakdowns and trigger maintenance activities. It also enables the subsequent evaluation of conducted maintenance and service activities by enabling an algorithm to scan reports (incl. free text fields) for key words which correspond to indicators for the success or failure of conducted activities.

3.3 Function tree

In literature, many characterizations of the term "prescriptive maintenance" and the structure of the corresponding activities are described, however the realization remains fuzzy. Only few sources go beyond simply pointing out the need for a "set of tasks" and these are not specific to onshore WTs [17,26]. So for this project a measure catalogue needed to be developed.

In order to develop a measure catalogue, a function tree is required to properly link recognized errors (e.g. from error messages) with faults in specific components. A function tree also extends to the sensor layer and their connections as an error can become observable on a level other than the underlying root cause. This is caused by the sensor setup in a WT. Thus, the function tree also enables the analysis of error codes regarding the "triggering" component [11]. An example would be the following error message: Error pitch system [Angle error blade A/B/C]. This can be assigned to the position sensor or the pitch motor in the function tree, for example. For this work, a function tree was developed in expert interviews with domain experts (such as wind farm operators or WT service engineers). The initial focus of the development was put on GE turbines (making up the majority of WTs in the wind farms under consideration), however the function trees can be expanded to WTs by other manufacturers in the future. As described above, the focus for the function tree was put on the pitch and converter system. In this function tree these systems and their sub-systems were structured. An extract is shown in Figure 4.

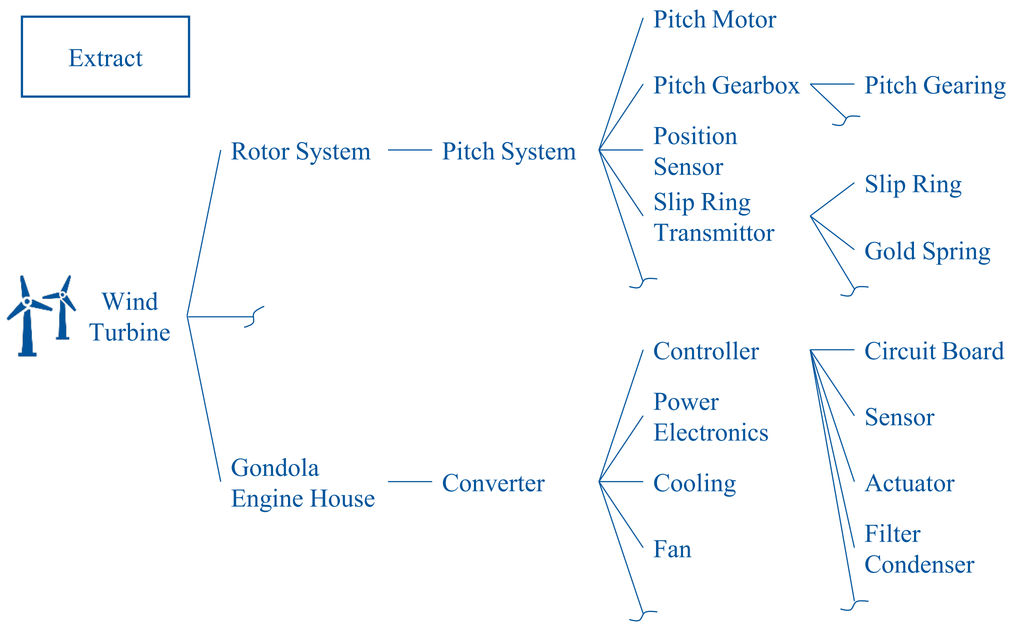


Figure 4: Function tree - pitch system and converter (Source: own representation)

3.4 Measure catalogue

Based on the developed function tree, interviews with domain experts to determine and validate specific actions for specific breakdowns on a component level were conducted. An initial set of maintenance measures was collected utilizing literature research and expert interviews. This set of measures was discussed, validated and extended. These measures are collected in the measure catalogue structured as shown in Table 1. When a certain error is recognized, a list of measures, which typically resolve the identified error is suggested to the maintenance technician.

Table 1: Excerpt from the developed measure catalogue (Source: own representation)

Category	Damage Mechanism	Maintenance Measure
Pitch system	Faulty battery	Change pitch battery
Pitch system	Position data outside tolerance	Check position sensor
Converter	Faulty converter controller	Check corresponding sensors, actuators and circuit board

The measures are currently mostly specific to GE WTs, but with moderate translation effort these can be applied to turbines of other makers as well. In addition to the expansion, an algorithm needs to be developed to connect the ontology with the measure catalogue to enable the automation of the prescriptive maintenance approach.

4. Summary and outlook

Wind power is a pillar of the renewable energy mix. The O&M costs for onshore wind turbines are a major addressable cost driver. The use of a prescriptive maintenance approach is a possibility to drive O&M costs down. A prescriptive maintenance approach utilizes insights from data analytics to derive maintenance activities and considers the object of maintenance activities to be a guiding and controlling element. However it heavily relies on (unstructured) data to be effective. An ontology can be used to make vast amounts of unstructured data (such as data from service reports) usable for such an approach, thus enabling SMEs without an abundancy of resources to pursue such approach.

This work demonstrated how this unstructured data can be used to carry out a prescriptive maintenance approach. First, based on an ontology, which enables automated analysis of event data of WTs, a function tree was developed to link errors with the underlying root-cause. Following this, measures to counter specific issues in specific components of WT systems and sub-systems were developed in close collaboration with domain experts. Then, these measures were collected in a measure catalogue, to provide maintenance and service technicians with the means to counter damage mechanisms with specific maintenance measures. The collection of measures and validation is ongoing. In a next step, the list of measures needs to be expanded, as the initial focus was only put on the pitch system and converter. Thus enabling a prescriptive maintenance approach, increasing maintenance effectiveness and lowering O&M costs.

In the next step, this measure catalogue needs to be expanded to other WT systems and WT manufacturers. In a further step an algorithm to identify relevant events in historical service reports will be developed. This algorithm will also provide users with an appropriate interface to fully realize the potentials of the developed approach. Also, further research with focus on monetary effects of the prescriptive maintenance approach and the corresponding maintenance activities is recommended to improve transparency of the return on maintenance (similar to the return on invest). This will allow for an improved and more holistic maintenance understanding and will unlock the potentials of a modern maintenance organization.

Acknowledgements

EUROPEAN UNION
Investing in our Future
European Regional
Development Fund

EFRE.NRW
Investitionen in Wachstum
und Beschäftigung

This project was funded by the European Regional Development Fund (EFRE) under the project number EU-2-2-029.

References

[1] Umweltbundesamt (Ed.), 2020, August 14. Windenergie. Retrieved from https://www.umweltbundesamt.de/themen/klima-energie/erneuerbare-energien/windenergie#strom. Accessed 13 April 2021.

[2] Morthorst, P. E., Auer, H., Garrad, A., Blanco, I., 2009. Wind Energy the Facts: The Economics of Wind Power. Retrieved from https://www.wind-energy-the-facts.org/images/chapter3.pdf. Accessed 2 April 2021.

[3] Umweltbundesamt (Ed.), 2021, March 4. Erneuerbare Energien in Zahlen. Retrieved from https://www.umweltbundesamt.de/themen/klima-energie/erneuerbare-energien/erneuerbare-energien-in-zahlen#uberblick. Accessed 13 April 2021.

[4] 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 6 April 2021.

[5] 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 21 March 2021.

[6] 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, 199–208. DOI:10.1016/J.RENENE.2017.03.097.

[7] Staffell, I., Green, R., 2014. How does wind farm performance decline with age? Renewable Energy 66, 775–786. DOI:10.1016/j.renene.2013.10.041.

[8] 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, 1261–1271. DOI:10.1016/j.rser.2017.09.100.

[9] Deutsches Institut für Normung (DIN) e. V. (Ed.), 2019, June. Grundlagen der Instandhaltung. Fundamentals of Maintenance. (Standard-No. DIN 31051:2019-06) Beuth, Berlin, 13 pp.

[10] IZP, IWES, SAG Erwin Peters GmbH, ENERTRAG AG, 2009. Erhöhung der Verfügbarkeit von Windkraftanlagen, EVW-Phase I. Teilprojekt: Entwicklung von zuverlässigkeitsbezogenen Betriebs- und Instandhaltungsstrategien für die Windenergie unter besonderer Berücksichtigung der Offshore-Bedingungen. EVW I, Förderkennzeichen 0327574, 132 pp. Retrieved from http://www.wind-evw.de/fileadmin/Dokumente/Abschlussberichte/EVW_I_Abschlussbericht/EVW_Phase_I_Abschlussbericht.pdf. Accessed 13 April 2021.

[11] Strack, B., Lenart, M., Frank, J., Kramer, N., 2021, March 29. Ontology for maintenance of onshore wind turbines. Forschung im Ingenieurwesen. Retrieved from https://www.springerprofessional.de/ontology-for-maintenance-of-onshore-wind-turbines/19017992. Accessed 13 April 2021.

[12] Schuh, G., Anderl, R., Dumitrescu, R., Krüger, A., Hompel, M. ten, 2020. Industry 4.0 Maturity Index: Managing the Digital Transformation of Companies – UPDATE 2020. acatech, Munich.

https://en.acatech.de/publication/industrie-4-0-maturity-index-update-2020/download-pdf?lang=en. Accessed 7 April 2021.

[13] Davenport, T. H., Harris, J. G., 2007. Competing on analytics: The new science of winning. Harvard Business School Press, Boston, Massachusetts.

[14] Deutsches Institut für Normung (DIN) e. V. (Ed.), 2018, February. Instandhaltung – Begriffe der Instandhaltung. Dreisprachige Fassung (EN 13306:2017). Maintenance – Maintenance terminology; Trilingual version (EN 13306:2017). (Standard-No. DIN EN 13306:2018-02) Beuth, Berlin.

[15] 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.

[16] Defèr, F., Birtel, F., Frank, J., Jussen, P., 2019. Smart Maintenance – just do it! Retrieved from https://www.fir.rwth-aachen.de/fileadmin/publikationen/whitepaper/fir-whitepaper-smart-maintenance-just-do-it.pdf. Accessed 13 April 2021.

[17] Clausen, U., Langkau, S., Kreuz, F. (Eds.), 2019. Advances in Production, Logistics and Traffic. Springer International Publishing, Cham.

[18] Faulstich, S., Hahn, B., 2019. Erkenntnisse aus digitalisierten Daten von über 3000 Windenergieanlagen im Betrieb, in: 10. VDI-Fachtagung Schwingungen von Windenergieanlagen 2019. Bremen, 4. und 5. Juni 2019. VDI-Berichte; Nr. 2346. VDI Verlag GmbH, Düsseldorf, pp. 1–16. Retrieved from http://publica.fraunhofer.de/eprints/urn_nbn_de_0011-n-5490812.pdf. Accessed 13 April 2021.

[19] Schuh, G., Peter, T., Gudergan, G., 2011. Designing Cooperation Concepts for Service Concepts, in: 2011 17th International Conference on Concurrent Enterprising. 20 – 22 June 2011, Aachen, Germany. IEEE, Piscataway, pp. 1–8.

[20] Parbs, H., 2017. Fragmentierung eines Service Ecosystem: Eine Grounded-Theory-Studie zur Instandhaltung in der Windenergiebranche. Bremen, Univ., Diss. Retrieved from https://media.suub.uni-bremen.de/bitstream/elib/1408/1/00106483-1.pdf. Accessed 13 April 2021.

[21] Jung, H., 2014. Verbundprojekt Erhöhung der Verfügbarkeit von Windenergieanlagen. EVW-Phase II: Praxisimplementierung von zuverlässigkeits- und zustandsorientierten Betriebs- und Instandhaltungsstrategien auf Basis von standardisierten Prozessabläufen. Forschungsvorhaben: 0327574D-H, 171 pp. Retrieved from https://www.wind-evw.de/fileadmin/Dokumente/Abschlussberichte/EVW_II_Abschlussbericht/EVW_Phase_II_Abschlussbericht_ohne_Anlagen.pdf. Accessed 13 April 2021.

[22] Studer, R., Benjamins, V.R., Fensel, D., 1998. Knowledge Engineering: Principles and Methods. Data & Knowledge Engineering (25), 161–197.

[23] Zhou, A., Yu, D., Zhang, W., 2014. A research on intelligent fault diagnosis of wind turbines based on ontology and FMECA. Advanced Engineering Informatics 29 (1), 115–125. DOI:10.1016/j.aei.2014.10.001.

[24] Papadopoulos, P., Cipcigan, L., 2009. Wind turbines' condition monitoring: an ontology model, in: 2009 International Conference on Sustainable Power Generation and Supply: SUPERGEN 2009. International Conference on Sustainable Power Generation and Supply. SUPERGEN 2009, Nanjing. 4/6/2009 – 4/7/2009. IEEE, pp. 1–4. DOI: 10.1109/SUPERGEN.2009.5430854.

[25] Ertek, G., Chi, X., Zhang, A.N., Asian, S., 2017. Text mining analysis of wind turbine accidents: An ontology-based framework, in: 2017 IEEE International Conference on Big Data (Big Data). 2017 IEEE International Conference on Big Data (Big Data 2017), Boston, MA. 12/11/2017 – 12/14/2017. IEEE, Piscataway, pp. 3233–3241. DOI: 10.1109/BigData.2017.8258305.

[26] Marques, H., Giacotto, A., 2019. Prescriptive Maintenance: Building Alternative Plans for Smart Operations, in: The 10th Aerospace Technology Congress, October 8 – 9, 2019, Stockholm, Sweden. 08.10.2019 – 28.09.2019. Linköping University Electronic Press, pp. 231–236. DOI:10.3384/ecp19162027.

Biography

Bernhard Strack is the head of the Competence Center for Maintenance (CCIH) and a Doctoral Candidate 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 Bachelor's and Master's degree in Mechanical Engineering and Business Administration from RWTH Aachen University.

Dr.-Ing. Jana Frank has been the head of the Service Management research department at the Institute for Industrial Management (FIR) at RWTH Aachen since 2019. Her work deals with the transformation of production companies – from providers of products to providers of integrated solutions. Jana Frank studied industrial engineering and materials engineering at RWTH Aachen University. She established and led the research area for subscription business management at FIR since 2011.

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.

Florian Pfau pursues his master's degree (M.Sc.) in Business Administration and Mechanical Engineering at RWTH Aachen University. He has been working at the Institute for Industrial Management (FIR) since 2020 as a research assistant with focus on maintenance operations and wind power, particularly in the ReStroK research project.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Machine Learning For Intelligent Maintenance And Quality Control: A Review Of Existing Datasets And Corresponding Use Cases

Nicolas Jourdan[1,*], Lukas Longard[1,*], Tobias Biegel[1,*], Joachim Metternich[1]
[1]TU Darmstadt, Institute of Production Management, Technology and Machine Tools (PTW), Darmstadt, Germany

Abstract

The advent of artificial intelligence and machine learning is influencing the manufacturing industry profoundly, enabling unprecedented opportunities to improve manufacturing processes within the three dimensions time, quality and cost. With the introduction of digitization and industry 4.0, increasing amounts of data become available for processing and use in smart manufacturing systems. However, the various use cases for machine learning in manufacturing often require problem-specific datasets for training and evaluation of algorithms which are difficult to acquire, hindering both practitioners and academic researchers in this area. As the respective data frequently contains sensitive information, manufacturing companies rarely release datasets to the public. Further, the relevant attributes and features of available datasets are usually not evident, requiring time-consuming analysis to evaluate if a dataset fits a given problem. As a result, it can be challenging to develop and evaluate machine learning methods for manufacturing systems due to the lack of an overview of available datasets. This paper presents a comprehensive overview of 47 existing, publicly available datasets, mapped to various use cases in manufacturing with the goal of simplifying and stimulating research. The characteristics of the datasets are compared using a set of descriptive attributes to provide an outline and guidance for further research and application of machine learning in manufacturing. In addition, suitable performance metrics for the evaluation of classification use cases in manufacturing are presented.

Keywords

Machine Learning; Artificial Intelligence; Manufacturing; Dataset; Benchmark; Metric; Evaluation;

1. Introduction and methodology

Machine Learning (ML) techniques increasingly transcend from research to practical applications in various industries. One of the industry areas that received significant attention in this context is the manufacturing industry [1]. The growing interest in manufacturing-related ML applications is fueled by the digitization of manufacturing processes in the context of industry 4.0 and the Internet of Things (IoT) [2]. However, the introduction of ML in manufacturing faces several challenges, with one of the most important being the acquisition of datasets for the development, training and evaluation of ML algorithms in high quantity. A sufficient data basis is crucial for the development of ML algorithms and strongly influences the achievable performance of the system [1]. Not only the quantity, but also the quality of the available data is of importance. Issues such as missing values, class imbalance, varying sampling frequencies and data types as well as high dimensionality have to be handled by the developers through pre-processing the data before

* Authors contributed equally

training the algorithms [3]. Compared to other ML application fields such as autonomous driving, only few publicly available datasets exist for manufacturing. Companies often see process data as sensitive information that cannot be shared due to privacy concerns [4]. As a result, the majority of studies that show successful applications of ML in manufacturing use cases do not share their training and testing datasets publicly, preventing an effective comparison between approaches [1]. In conclusion, an overview of publicly available datasets for ML applications in manufacturing is required to assist researchers and practitioners in the development and evaluation of algorithms and to enable the comparison of ML approaches in research studies. Few of such reviews (e.g. [4]) exist and to the best of the authors knowledge, none exist that account for modalities such as images, which are increasingly used in manufacturing-related ML applications [5–7]. In this study, the search for datasets was conducted on open platforms that provide dataset and code hosting for research and public competition purposes. The platforms include in alphabetical order: *GitHub [8], Kaggle [9], Mendeley Data [10], NASA Prognostics Center of Excellence (PCoE) [11], OpenML [12], University of California Irvine (UCI) Machine Learning Repository [13].* In addition to the resulting file storage ressources of the datasets, a search regarding accompanying publications that first release, describe and/or use the datasets for research, was conducted. Using the resulting publications, snowballing was applied to identify additional datasets that are hosted on platforms such as the universities of the corresponding authors. In total, 47 datasets have been identified and analysed regarding the comparison parameters. In addition to the selection or creation of an appropriate dataset for model training, the performance evaluation is an integral part of the model development process. Especially for classification tasks, the selection of an appropriate evaluation metric requires a deep understanding of the pursued task and relevant requirements [14]. Thus, a search for studies that utilize the identified datasets was conducted using Scopus, yielding 127 publications. These publications were consequently analysed regarding the applied performance evaluation metrics. The remainder of the paper is structured as follows: Chapter 2 introduces the identified datasets, sorted by the respective use cases. In Chapter 3, the most widespread classification metrics used for the evaluation of ML applications in manufacturing-related research are explained and critically analysed. Lastly, the conclusion is presented in Chapter 4.

2. Datasets and use cases in manufacturing

With the ongoing digitization there is an increasing number of research efforts emerging that focus on manufacturing-related problems. The datasets listed in this paper are suitable to address a subset of these problems and may be used to train and test ML methods specifically designed for manufacturing.

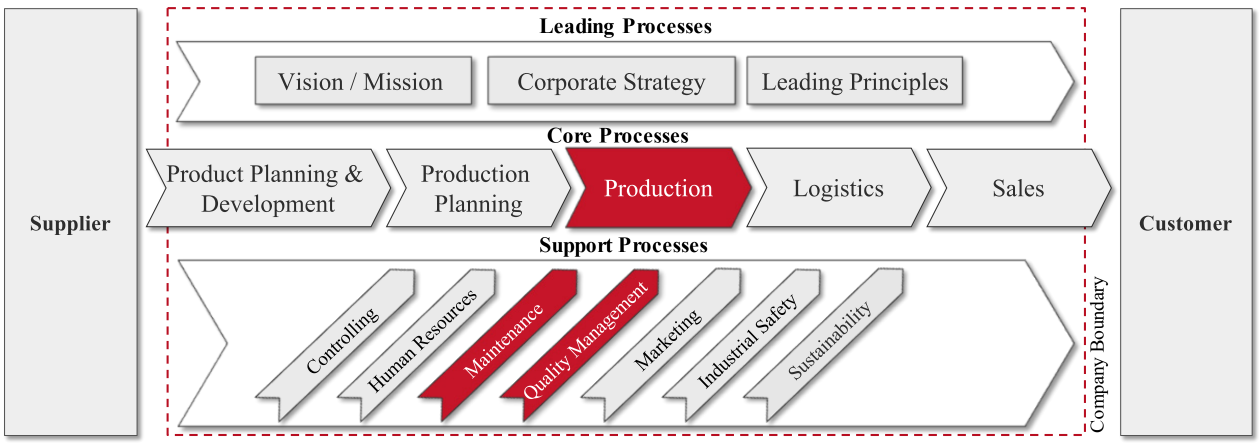


Figure 1: Process map highlighting the supporting areas that contain the use cases for all the presented datasets, adapted from [15].

In order to facilitate and further stimulate future research in this area, we identify common use cases for ML applications in manufacturing and map them to relevant datasets with the aim of accelerating the search for a suitable dataset based on the respective use case. The use cases defined in this paper are assigned to the two supporting processes: maintenance and quality management, as illustrated in the process map in Figure 1. The process map depicts the processes within the company boundaries that are necessary to meet customer demands. Each value adding activity requires support from indirectly value adding processes [15]. In terms of the core process production, the highlighted supporting processes embody two of the primary application areas of ML research in manufacturing nowadays and account for all the datasets listed in this paper [16].

The tables containing the identified datasets for the respective use cases are structured as follows: The *Name* column provides a short identifier to the datasets as well as a short description of the individual setting. The identifier is adopted from the original dataset source if available, otherwise it is newly created. Since a subset of the presented datasets is contained in [4], we adopted the respective names if applicable. Further, the *Type* and *Count* of the available non-target variable features is shown. In case of images, the resolution of the images is displayed as the feature count. The *Target Variable* column describes whether there are labels available for supervised learning: "C (*N*)" indicates that classification labels for *N* classes are given, while "R" indicates a supervised regression task. In some cases, labels for both, classification and regression, are available. The *Instances* column indicates the number of samples, e.g., rows, a given dataset contains. Further, *Official Train/Test Split* specifies, whether the dataset publishers provide a designated train / test data split for evaluation. This is especially important for comparability in research studies, as the same test split must be used to be able to compare the performance of different approaches. A consequent † indicates, that the target labels for the official test split are not publicly available, but rather hidden behind an evaluation server, which guarantees a fair benchmarking of approaches. The column *Data Source* highlights whether the dataset was collected from a real process, or rather generated synthetically using a simulation. Lastly, *Format* shows the file formatting of the raw data. In some cases, multiple formats are given, e.g., images in PNG-format and corresponding labels in XML-format. In the references section, the corresponding URL to the dataset is given together with the publication where the dataset was first introduced, if available. The use cases for each supporting process are presented in the following.

2.1 Maintenance (predictive maintenance and condition monitoring)

Predictive maintenance and condition monitoring are two terminologies that are used interchangeably by some researchers, while others view condition monitoring as part of the broader concept of predictive maintenance [17–20]. For the purpose of this paper, we follow the latter approach and thus consider condition monitoring as being a part of predictive maintenance. The use case predictive maintenance involves the data-driven assessment of the health status of machine components and sees its major objective in predicting the Remaining Useful Lifetime (RUL) of these components in order to reduce maintenance cost while simultaneously preventing unplanned downtimes. In machining applications such as milling, turning, or drilling, this, for instance involves the monitoring of the cutting tool to assess the current wear state, followed by the prediction of the RUL. Another typical scenario can be seen in the monitoring and RUL prediction of bearings. In this paper, we consider the data-driven health assessment as being integral to condition monitoring while the prediction of the RUL constitutes the broader case of predictive maintenance. The corresponding datasets for predictive maintenance and condition monitoring are exhibited in Table 1 on the following page.

Table 1: Datasets for predictive maintenance and condition monitoring

Name	Year	Features		Target Variable	Instances	Official Train/Test Split	Data Source	Format
		Type	Count					
Diesel Engine Faults Features [21] *Fault detection based on pressure curves and vibration.*	2020	Signal	84	C (4)	3.500	×	Syn.	MAT
Degradation of a Cutting Blade [22] *Wrapping machine process data over 12 months with a degrading cutting tool.*	2019	Signal	9	-	1.062.912	×	Real	CSV
CNC Mill Tool Wear [23] *CNC process data of wax milling with worn/unworn tools.*	2018	Signal	48	C (3*2)	25.286	×	Real	CSV
Condition Monitoring of Hydraulic Systems [24] *Test rig process data of multiple load cycles with various fault types and severity levels.*	2018	Signal	17	C (5*(2-4))	2.205	×	Real	Other
Production Plant Data for Condition Monitoring [22] *Anonymized process data of component run-to-failure experiments.*	2018	Signal	26	-	228.414	×	Real	CSV
Versatile Production System [25] *Popcorn production process data with multiple process steps.*	2018	Signal	5-85	-	80.000	×	Real	CSV
Degradation Measurement of Robot Arm Position Accuracy [26] *Target- and actual values of robotic arm tool position, velocity and current for health assessment.*	2017	Signal	73	-	155.000	×	Real	CSV
APS Failure at Scania Trucks [27] *Anonymized counters and histograms for air pressure system fault detection.*	2016	Signal	170	C (2)	76.000	✓	Real	CSV
Maintenance of Naval Propulsion Plants [28] *Gas turbine process data for component decay state prediction.*	2016	Signal	16	R	11.934	×	Syn.	Other
Plant Fault Detection [29] *Anonymized process data for plant fault detection.*	2015	Signal	10	C (6)	8.938.370	×	Real	CSV
Asset Failure and Replacement [30] *Anonymized data for asset fault detection.*	2014	Signal	1	C (2)	447.341	✓†	Real	CSV
Maintenance Action Recommendation [31] *Anonymized process and maintenance data of an industrial asset for maintenance action recommendation.*	2013	Signal	32	C (14)	2.097.152	✓†	Real	CSV
Anemometer Fault Detection [32] *Anemometer measurements for fault detection*	2011	Signal	16 16-20	-	345.700 208.800	✓†	Real	Other
Gearbox Fault Detection [33] *Test rig accelerometer data for fault detection.*	2009	Signal	3	-	> 10 Mio.	×	Real	CSV
Li-Ion Battery Aging [34] *Battery test rig data during charge and discharge cycles for degradation detection.*	2008	Signal	12	-	2.167	×	Real	MAT
Turbofan Engine Degradation Simulation [35] *C-MAPSS simulation sensor data of various conditions and fault modes.*	2008	Signal	26	-	262.256	✓	Syn.	Other
Bearing [36] *Bearing test rig accelerometer data of run-to-failure experiments.*	2007	Signal	4-8	-	61.440	×	Real	CSV
Milling [37] *Milling process- and external sensor data for tool wear detection.*	2007	Signal	13	R	1.503.000	×	Real	MAT
CWRU Bearing Data [38] *Bearing test rig accelerometer data for fault detection.*	n.A.	Signal	5	C (2)	> 10 Mio.	×	Real	MAT

2.2 Quality management

The subprocess quality management embodies the use cases process monitoring, predictive quality, quality inspection and process parameter optimization. The respective use cases and corresponding datasets will be introduced in the following subsections.

2.2.1 Process monitoring

The analysis of sensor-based process data can yield valuable information for the purpose of process control and quality monitoring [39]. The idea of process monitoring is to understand the variation in a process and to assess its current state [40]. A widely used technique in this field is control charting which involves two distinct monitoring phases, i.e. phase I and phase II [41]. In phase I, control charts are used to retrospectively test whether the process was in control after the data have been sampled from the process. The result of this phase is a Normal Operating Condition (NOC) dataset in which the underlying process is assumed to be in-control. With the help of the NOC dataset, control limits are established based on which a new observation of process data will be evaluated. This is the objective of phase II. Process monitoring has been an active research field throughout the last decades [42]. Especially within the process industry, the application of Multivariate Statistical Process Monitoring (MSPM) methods gained popularity [43,44]. In terms of discrete manufacturing, recent research focusses on the initiation of a paradigm shift from the conventional post-process Statistical Process Control (SPC), i.e. inferring the process condition based on measurements taken from the manufactured product to the so called in-process SPC that aims at inferring the process condition based on actual process data [45]. In both fields of industry, the application of ML, especially Deep Learning (DL) is receiving more and more attention and provides promising results for future research in this field [46,47]. The corresponding datasets for process monitoring are exhibited in Table 2.

Table 2: Datasets for process monitoring

Name	Year	Features		Target Variable	Instances	Official Train/Test Split	Data Source	Format
		Type	Count					
High Storage System Anomaly Detection [48] *Storage test rig process data for anomaly detection.*	2018	Signal	20	C (2)	91.000	×	Syn.	CSV
Genesis Pick-and-Place Demonstrator [49] *Material sorting test rig process data for anomaly detection.*	2018	Signal	23	C (3)	32.440	×	Real	CSV
Tennessee Eastman Process Simulation Dataset [50] *Simulated chemical process data for anomaly detection with different fault types.*	2017	Signal	51	C (21) / R	> 10 Mio.	✓	Syn.	RData
Robot Execution Failures [51] *Force and torque measurements of an industrial robot with different erroneous operating conditions.*	1999	Signal	89	C (13)	463	×	Real	Other
Mechanical Analysis [52] *Vibration measurements of electromechanical devices with different erroneous operating conditions.*	1990	Signal	7	C (6)	209	✓	Real	MAT
CWRU Bearing Data [38] *Bearing test rig accelerometer data for anomaly detection.*	n/a	Signal	5	C (2)	> 10 Mio.	×	Real	MAT

2.2.2 Predictive quality and quality inspection

The use case predictive quality incorporates the scenario where the prediction of the product quality is of primary concern. The accurate prediction of the product quality can be used to better control the manufacturing process [53]. The costs of delayed discovery of nonconformities in the product lifecycle increase exponentially the further the product moves down the value-chain [54]. Therefore, it becomes useful to predict if a product will fail specification tests in later stages of the process if the cycle times of a process chain are very long [55].

Table 3: Datasets for predictive quality and quality inspection.

Name	Year	Features		Target Variable	Instances	Official Train/Test Split	Data Source	Format
		Type	Count					
Casting Product Quality Inspection [6] *Grayscale images of pump impeller castings with and without defects.*	2020	Image	300×300 512×512	C (2)	7.348	✓	Real	JPG
GC10-DET [56] *Grayscale images of metal surfaces with various defect types and corresponding bounding box annotations.*	2020	Image	Varying	C (10)	3.570	×	Real	JPG, XML
Mechanic Component Images [7] *Grayscale images of air conditioner pistons with various defect types.*	2020	Image	86×90	C (3)	285	×	Real	PNG
Multi-Stage Continuous Flow Process [57] *Anonymized process data of a production line with quality measurements of part dimensions.*	2020	Signal	116	-	14.088	×	Real	CSV
Plastic Extrusion Defects [58] *Process data of a plastic extrusion process.*	2020	Signal	470	-	226.536	×	Real	CSV
AITEX [59] *Grayscale images of textile fabrics with various defect types and corresponding segmentation masks.*	2019	Image	4096×256	C (13)	245	×	Real	PNG, Mask
Deep PCB [60] *Grayscale images of circuit boards with various defect types and corresponding bounding box annotations.*	2019	Image	640×640	C (7)	1.500	✓	Real	JPG, Mask
Severstal Steel Defect Detection [61] *Grayscale images of steel surfaces with various defect types and corresponding segmentation polygons.*	2019	Image	1600×256	C (5)	18.074	✓†	Real	JPG, CSV
Turning Dataset for Chatter Diagnosis [62] *Sensory data of a turning test rig and varying strengths of chatter.*	2019	Signal	8	C (4)	>10 Mio.	×	Real	MAT
Magnetic Tile Defect [63] *Grayscale images of magnetic tile surfaces with various defect types and corresponding segmentation masks.*	2018	Image	248×373	C (6)	1.344	×	Real	JPG, PNG
TIG Welding [5] *Grayscale images of a welding process with various defect types.*	2018	Image	800×974	C (6)	33.254	✓	Real	PNG, JSON
Mining Process [64] *Process data of a mining process for impurity prediction in ore concentrate.*	2017	Signal	24	R	737.454	×	Real	CSV
Bosch Production Line Performance [65] *Anonymized process data of production lines with and without defects.*	2016	Signal	4264	C (2)	2.368.43 5	✓†	Real	CSV
WM811K Wafer Maps [66] *Defect matrices of semiconductor wafers with various defect types.*	2014	2D Defect Matrix	Varying	C (9)	811.457	×	Real	MAT
NEU Surface Defect Database [67] *Grayscale images of metal surfaces with various defect types and corresponding bounding box annotations.*	2013	Image	200×200	C (6)	1.800	×	Real	BMP, XML
Steel Plate Faults [68] *Geometric measurements of steel plates with various defect types.*	2010	Signal	27	C (7)	1.941	×	Real	CSV
HCI Industrial Optical Inspection [69] *Synthetic grayscale images of textured surfaces with corresponding defect ellipses.*	2007	Image	512×512	C (2)	16.100	✓	Syn.	PNG, Other

In practice, the application of predictive quality requires the existence of sufficient quality data to find the dependencies between the generally more accessible process data on the basis of which the quality of the product shall be predicted in the future. This can be difficult especially in terms of low volume discrete production systems [70]. Typical applications of predictive quality can be seen in the prediction of the surface quality, surface roughness as well as deformations or chatter marks [71]. In this paper, quality inspection entails the assessment of the quality of a manufactured product at certain stages of the manufacturing process. In a recent review paper [72] the authors conducted a thorough investigation based on the last three decades of the state of the art in so called zero defect manufacturing. The authors subdivide quality inspection based on the respective manufacturing stage into three different phases, i.e. prior to, during or after the manufacturing of the product. In terms of this study, we summarize all three aforementioned phases under the term quality inspection. ML Methods such as Support Vector Machine (SVM), Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are used in this field for signal- and image processing with the goal of assessing the quality of the manufactured parts [5]. The corresponding datasets are exhibited in Table 3 on the previous page.

2.2.3 Process parameter optimization

Process parameters are generally chosen based on human judgement and experience in combination with the use of handbooks that provide recommendations, which may lead to a loss of productivity and quality [73]. Consequently, the selection of the optimal process parameters such as cutting speed, depth of cut, etc. plays an important role in today's highly competitive manufacturing industries and provides the opportunity to achieve high quality products with less cost and time constraints [74]. The field of application for process parameter optimization with the help of ML has received a lot of interest in recent research. In [75] the authors provide an extensive review for the application of ML for the optimization of process parameters. The main areas mentioned include milling, turning, gear hobbing and boring, finishing, welding and plastic injection molding. Next to supervised ML methods such as ANN or SVM, evolutionary optimization techniques such as Genetic Algorithms (GA), Particle Swarm Optimization (PSO) or Simulated Annealing (SA) have been used for process parameter optimization [76]. The corresponding datasets are exhibited in Table 4.

Table 4: Datasets for process parameter optimization.

Name	Year	Features		Instances	Official Train/Test Split	Data Source	Format
		Type	Count				
Laser Welding [77] *Process parameter recordings for correlation with weld quality indicators such as weld depth and geometrical dimensions.*	2020	Signal	13	361	×	Real	XLS
3D Printer [78] *Process parameters of a 3D printer for correlation with print quality indicators such as roughness, tension and elongation.*	2018	Signal	12	50	×	Real	CSV
Tool Path Generation [79] *Shape deviation measurements and corresponding simulated cutting conditions.*	2018	Signal	9	4.968	×	Real, Syn.	CSV
Mercedes-Benz Greener Manufacturing [80] *Car feature configurations to be correlated with the required test time of the configurations.*	2017	Signal	378	8.420	✓†	Real	CSV
SECOM [81] *Semiconductor process measurements and corresponding yields for determination of key factors to yield.*	2008	Signal	591	1.567	×	Real	Other

3. Evaluation metrics

Besides data selection, another integral part of the model development process in ML is the performance evaluation. The aim of the performance evaluation is to find a model that best represents the underlying data and also performs well on new data [1]. For this purpose, a model is evaluated on a separate held-out test set using appropriate performance metrics after the training process. The selection of an appropriate evaluation metric requires a deep understanding of the pursued task with all its characteristics [82]. Despite the many discussions in the field of performance metrics in science, misleading or inadequate metrics are often used [83]. The majority of the identified datasets in this paper are suitable for classification as well as regression tasks. Both types require task-specific metrics to evaluate the respective performance of the models. Due to the large number of available metrics and their susceptibility to changing framework conditions (e.g. imbalanced classes), the selection of classification metrics often turns out to be difficult. Haixiang et al. evaluated 517 papers concerned with imbalanced classification across multiple domains and found out that 201 out of those (38%) were using accuracy as an evaluation metric [84]. In contrast, for regression, the relation and appropriateness of several evaluation metrics have been analysed thoroughly [85,86] and the difference between existing metrics is sufficiently clear. Moreover through the continuous character of the output (and measures), the selection of metrics is facilitated [14]. As a result, only classification metrics are further elaborated and critically discussed in the context of manufacturing in the following. Subsequently, the distribution of these is evaluated by analysing papers related to the datasets found.

3.1 Classification metrics

A common method for evaluating the performance of classifiers is the confusion matrix. It is applicable for problems where the output includes two or more classes. In the confusion matrix for binary classification problems, the classes are called positive and negative while the labels true and false indicate whether a prediction matches the true value or not. Most of the classification performance metrics can be derived directly (e.g. sensitivity, precision) or indirectly (e.g. Receiver Operating Characteristic (ROC), Precision-Recall Curve (PRC)) from the confusion matrix [83]. Accuracy describes the portion of correctly predicted data points out of all data points. While it is often used as a single metric to evaluate classification problems, the pure focus on maximizing accuracy is viewed critically by some researchers [83,87]. The reason for this is that classification accuracy considers the same misclassification costs for false positive and false negative errors. For most real-world problems one type of classification error (i.e. type I, type II) is more expensive than another. This issue is especially important when dealing with imbalanced datasets which frequently appear in manufacturing use cases. Suppose a model predicts NOK parts (positive class) at a quality gate which represents a problem with two classes: class A (OK parts) is 95% of the dataset and class B (NOK parts) is the remaining 5%. By simply predicting class A for every sample, the model can reach an accuracy of 95%, which seems to be a good score, but it is not. To overcome this, Seliya et al. point out that a classifier should be evaluated not only by one, but a set of performance metrics. Through this approach, several performance aspects can be considered and differentiated conclusions can be drawn [88,89]. In the given example of quality control, the correct prediction of the minority class may be of higher importance since a faulty delivery to the customer is to be avoided at all costs which promotes the use of recall as the primary evalution metric. Though, precision cannot be ignored as a low precision may lead to high quality control costs due to a high number of tests. This highlights, that the selection of a relevant metric is highly dependend on the actual use case. A metric that takes both recall and precision into account is the F-score. It uses the harmonic mean in place of the arithmetic mean, thus punishing the extreme values more [82]. A special case of this metric is the F_β-score, that allows the user to emphasize on either recall or precision [90]. Neither of the above mentioned metrics take into account the number of true negatives [91]. Specificity is used to determine the proportion of actual negative cases which got predicted correctly.

All metrics mentioned so far are single-threshold metrics, which means that they are defined for an individual score threshold (cut-off) of a classifier and cannot give an overview of the different performance levels at varying thresholds [90]. Through performance curves, the changing metrics at varying thresholds can be captured [90]. The most widespread curves are the ROC and the PRC. The ROC plot shows the trade-off between recall and specificity at varying thresholds [92] and an operating point, i.e. threshold, needs to be chosen according to the use case requirements. A single performance metric that can be derived from the ROC curve is the Area Under the ROC Curve (AUROC) score. An AUROC of 0.5 results from random choice while an AUROC of 1.0 shows a perfect classifier [93]. As with the ROC curve, the Area Under the PRC (AUPRC) is also used as a single metric. Differently to AUROC though, the baseline of AUPRC changes with class imbalance [90].

3.2 Use of classification metrics in publications

After presenting and discussing the state of the art in terms of classification metrics and the associated difficulties, the analysis of 49 different publications dealing with classification algorithms on the identified datasets is explained below. The selection is based on a backward search starting from the datasets found. Similar to the findings of Haixiang et al., accuracy is the most widely used metric in classification tasks. Almost 72% (35) of all analysed publications use accuracy as a performance metric, while for 39% of the publications, accuracy was the only metric used (see Figure 2). Furthermore, 45% of the publications only use one metric to evaluate their results. Although several authors highlight the widespread use of performance curves as an evaluation metric, this could not be fully confirmed in the analysis conducted. Altogether only eleven publications either used ROC, AUROC, PRC or AUPRC to evaluate their results. It should be noted that none of these used both ROC (AUROC) or PRC (AUPRC) and thus could not encompass all performance aspects. Only about 25% (12) of the publications studied used three or more metrics for evaluation.

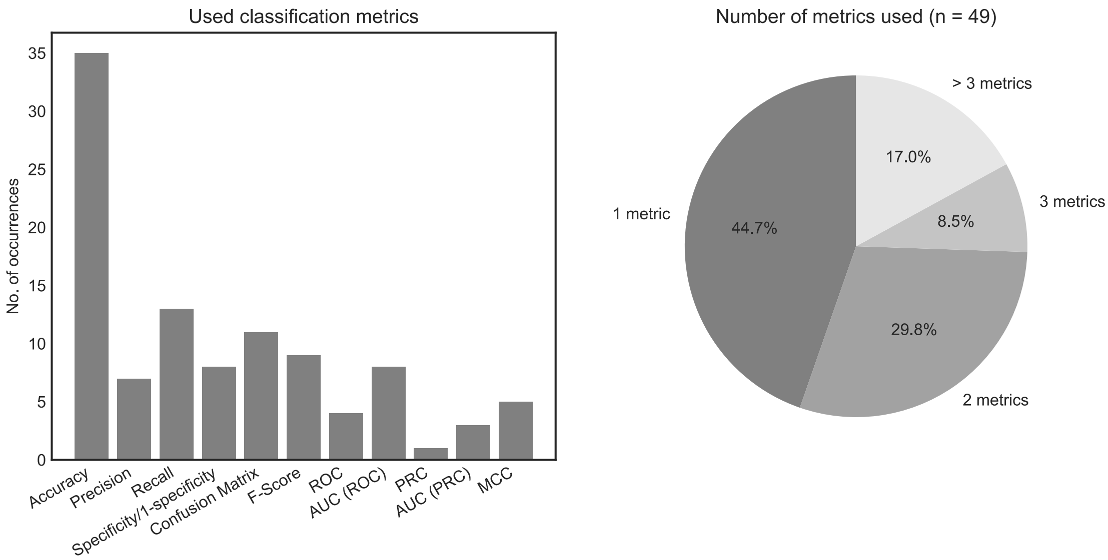


Figure 2: Analysis of publications regarding evaluation metrics (n = 49)

4. Conclusion

In this paper we provide a comprehensive overview and comparison of datasets suitable for the development of ML applications in the manufacturing sector as well as corresponding metrics for effective performance evaluation of classification problems. The identified use cases include predictive maintenance, condition monitoring, process monitoring, predictive quality, quality inspection and process parameter optimization. The analysis has the aim of stimulating research in this field as well as to promote the use of public datasets for evaluation. This is required to compare the performance of different approaches objectively in research, which is often not possible due to the use of proprietary datasets that are not shared because of data privacy concerns. Further, manufacturing companies can employ the public datasets in the development of algorithms for their specific facilities and gain practical knowledge and experience in the process [4]. Additionally, we showed that a large part of the analysed studies solely use accuracy for performance evaluation of classification problems, which may not be expressive enough in all use cases. As an opportunity for future work, it would be of interest to identify or create datasets, corresponding tasks and metrics that can serve as a standard benchmark for certain use cases in manufacturing, comparable to other industry areas such as autonomous driving.

Acknowledgements

The research leading to these results has received funding from the European Institute of Technology (EIT Manufacturing) under grant agreement number 21026 (IVE), as well as from the German Federal Minisitry of Education and Research under grant agreement number 02L19C150 (KompAKI).

References

[1] 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.

[2] Diez-Olivan, A., Del Ser, J., Galar, D., Sierra, B., 2019. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion 50, 92–111. https://www.sciencedirect.com/science/article/pii/S1566253518304706.

[3] Pham, D.T., Afify, A.A., 2005. Machine-learning techniques and their applications in manufacturing. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture 219 (5), 395–412.

[4] Krauß, J., Dorißen, J., Mende, H., Frye, M., Schmitt, R.H., 2019. Machine learning and artificial intelligence in production: application areas and publicly available data sets. Production at the leading edge of technology, 493–501.

[5] Bacioiu, D., Melton, G., Papaelias, M., Shaw, R., 2019. Automated defect classification of Aluminium 5083 TIG welding using HDR camera and neural networks. Journal of Manufacturing Processes 45, 603–613. https://www.kaggle.com/danielbacioiu/tig-aluminium-5083. Accessed 31 March 2021.

[6] Kantesaria, N., Vaghasia, P., Hirpara, J., Bhoraniya, R., 2020. Casting product image data for quality inspection. https://www.kaggle.com/ravirajsinh45/real-life-industrial-dataset-of-casting-product. Accessed 19 March 2021.

[7] Paladi, S., 2020. Mechanic Component Images. https://www.kaggle.com/satishpaladi11/mechanic-component-images-normal-defected. Accessed 19 March 2021.

[8] GitHub. https://github.com/. Accessed 31 March 2021.

[9] Kaggle. https://www.kaggle.com/. Accessed 31 March 2021.

[10] Mendeley Data. https://data.mendeley.com/. Accessed 31 March 2021.

[11] NASA Ames Prognostics Center of Excellence (PCoE);. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/. Accessed 27 November 2020.

[12] OpenML. https://www.openml.org/home. Accessed 31 March 2021.

[13] University of California Irvine. UCI Machine Learning Repository. University of California Irvine. https://archive.ics.uci.edu. Accessed 3 November 2020.

[14] Ferri, C., Hernández-Orallo, J., Modroiu, R., 2009. An experimental comparison of performance measures for classification. Pattern Recognition Letters 30 (1), 27–38.

[15] Bogaschewsky, R., Rollberg, R., 1998. Prozeßorientiertes Management. Springer, Berlin, Heidelberg.

[16] Kang, Z., Catal, C., Tekinerdogan, B., 2020. Machine learning applications in production lines: A systematic literature review. Computers & Industrial Engineering 149, 106773.

[17] Goyal, D., Saini, A., Dhami, S.S., Pabla, B.S., 2016. Intelligent predictive maintenance of dynamic systems using condition monitoring and signal processing techniques — A review. 2016 International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Spring), 1–6.

[18] Wang, H., Ye, X., Yin, M., 2016. Study on predictive maintenance strategy. International Journal of u-and e-Service, Science and Technology (9), 295–300.

[19] Hu, J., Chen, P., 2020. Predictive maintenance of systems subject to hard failure based on proportional hazards model. Reliability Engineering & System Safety 196.

[20] Mobley, R.K., 2002. An introduction to predictive maintenance, 2nd ed. Butterworth-Heinemann, Amsterdam, New York, 438 pp.

[21] Pestana, D., 2020. Diesel engine faults features dataset: (3500-DEFault). https://data.mendeley.com/datasets/k22zxz29kr/1. Accessed 20 March 2021.

[22] Birgelen, A.v., Buratti, D., Mager, J., Niggemann, O., 2018. Self-organizing maps for anomaly localization and predictive maintenance in cyber-physical production systems. Procedia CIRP 72, 480–485. https://www.kaggle.com/inIT-OWL/one-year-industrial-component-degradation; https://www.kaggle.com/inIT-OWL/production-plant-data-for-condition-monitoring. Accessed 31 March 2021.

[23] Kovalenko, I., Saez, M., Barton, K., Tilbury, D., 2017. SMART: A system-level manufacturing and automation research testbed. Smart Sustain. Manuf. Syst. 1 (1), 20170006. https://www.kaggle.com/shasun/tool-wear-detection-in-cnc-mill. Accessed 31 March 2021.

[24] Helwig, N., Pignanelli, E., Schutze, A., 2015. Condition monitoring of a complex hydraulic system using multivariate statistics. IEEE International Instrumentation and Measurement Technology Conference, 210–215. https://archive.ics.uci.edu/ml/datasets/Condition+monitoring+of+hydraulic+systems. Accessed 31 March 2021.

[25] Institut für industrielle Informationstechnik - inIT, Technische Hochschule Ostwestfalen-Lippe, 2018. Versatile Production System. https://www.kaggle.com/inIT-OWL/versatileproductionsystem. Accessed 20 March 2021.

[26] Qiao, H., 2017. Degradation measurement of robot arm position accuracy. National Institute of Technology (NIST). https://www.nist.gov/el/intelligent-systems-division-73500/degradation-measurement-robot-arm-position-accuracy. Accessed 19 March 2021.

[27] Lindgren, T., Biteus, J., 2016. APS failure at Scania trucks data set. Scania CV AB. https://archive.ics.uci.edu/ml/datasets/APS+Failure+at+Scania+Trucks. Accessed 20 March 2021.

[28] Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., Figari, M., 2016. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. Proceedings of the IMechE 230 (1), 136–153. http://archive.ics.uci.edu/ml/datasets/Condition+Based+Maintenance+of+Naval+Propulsion+Plants. Accessed 31 March 2021.

[29] PHM Society. 2015 PHM Society Conference Data Challenge. PHM Society. https://github.com/robot007/PHM15. Accessed 31 March 2021.

[30] PHM Society, 2014. 2014 PHM Society Conference Data Challenge. PHM Society. https://phmsociety.org/conference/annual-conference-of-the-phm-society/annual-conference-of-the-prognostics-and-health-management-society-2014/phm-data-challenge-2/. Accessed 18 March 2021.

[31] PHM Society, 2013. 2013 PHM Society Conference Data Challenge. PHM Society. https://phmsociety.org/conference/annual-conference-of-the-phm-society/annual-conference-of-the-prognostics-and-health-management-society-2013/phm-data-challenge/. Accessed 31 March 2021.

[32] PHM Society, 2011. 2011 PHM Society Conference Data Challenge. PHM Society. https://phmsociety.org/phm_competition/2011-phm-society-conference-data-challenge/. Accessed 18 March 2021.

[33] Goebel, K., 2009. Gearbox Fault Detection Dataset, PHM Data Challenge 2009. https://c3.nasa.gov/dashlink/resources/997/. Accessed 31 March 2021.

[34] Mcintosh, D., 2010. Li-ion battery aging datasets. NASA Ames Prognostics Center of Excellence (PCoE). https://c3.nasa.gov/dashlink/resources/133/. Accessed 20 March 2021.

[35] Saxena, A., Goebel, K., 2008. Turbofan engine degradation simulation data set. NASA Ames Prognostics Center of Excellence (PCoE). https://c3.nasa.gov/dashlink/resources/139/. Accessed 20 March 2021.

[36] Lee, J and Qiu, H and Yu, G and Lin, Ja and others, 2007. Bearing data set. IMS, University of Cincinnati, NASA Ames Prognostics Center of Excellence (PCoE), Rexnord Technical Services. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/. Accessed 20 March 2021.

[37] Agogino, A., Goebel, K., 2007. Milling data set. NASA Ames Prognostics Data Repository; BEST lab, UC Berkeley. https://ti.arc.nasa.gov/tech/dash/groups/pcoe/prognostic-data-repository/. Accessed 20 March 2021.

[38] Case Western Reserve University. CWRU Bearing Data Center. Case Western Reserve University. https://csegroups.case.edu/bearingdatacenter. Accessed 19 March 2021.

[39] D. E. Lee, Inkil Hwang, C. M. O. Valente, J. F. G. Oliveira, David A. Dornfeld, 2006. Precision manufacturing process monitoring with acoustic emission. Condition monitoring and control for intelligent manufacturing, 33–54.

[40] William H. Woodall, Douglas C. Montgomery, 2014. Some current directions in the theory and application of statistical process monitoring. Journal of Quality Technology 46 (1), 78–94.

[41] Bersimis, S., Panaretos, J., Psarakis, S., 2005. Multivariate statistical process control charts and the problem of interpretation: a short overview and some applications in industry, in: Proceedings of the 7th Hellenic European Conference on Computer Mathematics and its Applications, Athens Greece.

[42] Tang, P., Peng, K., Dong, J., Zhang, K., Zhao, S., 2020. Monitoring of nonlinear processes with multiple operating modes through a novel gaussian mixture variational autoencoder model. IEEE Access 8, 114487–114500.

[43] Qin, S.J., 2003. Statistical process monitoring: basics and beyond. J. Chemometrics 17 (8-9), 480–502.

[44] Yu, J., Liu, X., Ye, L., 2021. Convolutional long short-term memory autoencoder-based feature learning for fault detection in industrial processes. IEEE Trans. Instrum. Meas. 70, 1–15.

[45] Maggioni, M., Marzorati, E., Grasso, M., Colosimo, B.M., Parenti, P., 2014. In-process quality characterization of grinding processes: A sensor-fusion based approach. ASME 2014 12th Biennial Conference.

[46] Lv, F., Fan, X., Wen, C., Bao, Z., 2018. Stacked sparse auto encoder network based multimode process monitoring. 2018 International Conference on Control, Automation and Information Sciences (Iccais), 227–232.

[47] Yan, S., Yan, X., 2020. Quality-driven autoencoder for nonlinear quality-related and process-related fault detection based on least-squares regularization and enhanced statistics. Ind. Eng. Chem. Res. 59 (26), 12136–12143.

[48] Hranisavljevic, N., Niggemann, O., Maier, A. A novel anomaly detection algorithm for hybrid production systems based on deep learning and timed automata. arXiv preprint arXiv:2010.15415 2020. https://www.kaggle.com/inIT-OWL/high-storage-system-data-for-energy-optimization. Accessed 31 March 2021.

[49] Birgelen, A.v., Niggemann, O., 2018. Anomaly detection and localization for cyber-physical production systems with self-organizing maps. IMPROVE-Innovative Modelling Approaches for Production Systems to Raise

Validatable Efficiency, 55–71. https://www.kaggle.com/inIT-OWL/genesis-demonstrator-data-for-machine-learning/home. Accessed 31 March 2021.

[50] Rieth, C.A., Amsel, B.D., Tran, R., Cook, M.B., 2017. Additional tennessee eastman process simulation data for anomaly detection evaluation. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/6C3JR1. Accessed 20 March 2021.

[51] Camarinha-Matos, L.M., Lopes, L.S., Barata, J., 1996. Integration and learning in supervision of flexible assembly systems. IEEE Trans. Robot. Automat. 12 (2), 202–219. http://archive.ics.uci.edu/ml/datasets/Robot+Execution+Failures. Accessed 31 March 2021.

[52] Bergadano, F., Giordana, A., Saitta, L., Marchi, D. de, Brancadori, F., 1990. Integrated learning in a real domain. Machine Learning Proceedings, 322–329. http://archive.ics.uci.edu/ml/datasets/Mechanical+Analysis. Accessed 31 March 2021.

[53] Xu, P., Ren, Z., Shen, Y., Yu, W., He, L., 2021. Quality prediction of discrete manufacturing process based on CGAN&Catboost hybrid model. J. Phys.: Conf. Ser. 1757 (1), 12072.

[54] Jimmy Chhor, Stefan Gerdhenrichs, Felix Mohrschladt, Robert H. Schmitt, 2019. Development of a machine learning model for a multi-correlative sample-based prediction of product quality for complex machining processes, in: , Production at the leading edge of technology. Springer Vieweg, Berlin, Heidelberg, pp. 523–532.

[55] Krauß, J., Pacheco, B.M., Zang, H.M., Schmitt, R.H., 2020. Automated machine learning for predictive quality in production. Procedia CIRP 93, 443–448. https://www.sciencedirect.com/science/article/pii/s2212827120306016.

[56] Lv, X., Duan, F., Fu, X., Gan, L., 2020. Deep metallic surface defect detection: The new benchmark and detection network. Sensors. https://www.kaggle.com/zhangyunsheng/defects-class-and-location.

[57] Liveline Technologies, 2020. Multi-stage continuous-flow manufacturing process. Liveline Technologies. https://www.kaggle.com/supergus/multistage-continuousflow-manufacturing-process/metadata. Accessed 19 March 2021.

[58] Kaggle, 2020. Find a defect in the production extrusion line. https://www.kaggle.com/podsyp/find-a-defect-in-the-production-extrusion-line/metadata. Accessed 20 March 2021.

[59] Silvestre-Blanes, J., Albero-Albero, T., Miralles, I., Pérez-Llorens, R., Moreno, J., 2019. A public fabric database for defect detection methods and results. Autex Research Journal (4), 363–374. https://www.aitex.es/afid/. Accessed 31 March 2021.

[60] Tang, S., He, F., Huang, X., Yang, J. Online PCB defect detector on a new PCB defect dataset. https://github.com/Charmve/Surface-Defect-Detection/tree/master/DeepPCB. Accessed 31 March 2021.

[61] Severstal, 2019. Steel defect detection. Severstal. https://www.kaggle.com/c/severstal-steel-defect-detection/overview. Accessed 19 March 2021.

[62] Yesilli, M., 2019. Turning dataset for chatter diagnosis using machine learning. Mendeley Data. http://dx.doi.org/10.17632/hvm4wh3jzx.1.

[63] Huang, Y., Qiu, C., Guo, Y., Wang, X., Yuan, K., 2020. Surface defect saliency of magnetic tile. The Visual Computer (36), 85--96. https://github.com/abin24/Magnetic-tile-defect-datasets. Accessed 31 March 2021.

[64] Magalhães Oliveira, E., 2017. Quality prediction in a mining process. https://www.kaggle.com/edumagalhaes/quality-prediction-in-a-mining-process. Accessed 19 March 2021.

[65] Robert Bosch GmbH, 2016. Bosch production line performance. Robert Bosch GmbH. https://www.kaggle.com/c/bosch-production-line-performance/overview. Accessed 19 March 2021.

[66] National Taiwan University, CS Dept. WM811K: Wafer Map. National Taiwan University, CS Dept. http://mirlab.org/dataSet/public/. Accessed 19 March 2021.

[67] Song, K., Yan, Y., 2013. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects. Applied Surface Science 285, 858–864. http://faculty.neu.edu.cn/yunhyan/NEU_surface_defect_database.html. Accessed 31 March 2021.

[68] Semeion, Research Center of Sciences of Communication, Via Sersale 117, 00128, Rome, Italy. Steel Plate Faults Dataset. https://archive.ics.uci.edu/ml/datasets/steel+plates+faults. Accessed 11 May 2021.

[69] Wieler, M., Hahn, T., 2007. Weakly supervised learning for industrial optical inspection. DAGM. https://hci.iwr.uni-heidelberg.de/content/weakly-supervised-learning-industrial-optical-inspection. Accessed 19 March 2021.

[70] Gittler, T., Relea, E., Corti, D., Corani, G., Weiss, L., Cannizzaro, D., Wegener, K., 2019. Towards predictive quality management in assembly systems with low quality low quantity data – a methodological approach. Procedia CIRP 79, 125–130.

[71] Kim, D.-H., Kim, T.J.Y., Wang, X., Kim, M., Quan, Y.-J., Oh, J.W., Min, S.-H., Kim, H., Bhandari, B., Yang, I., Ahn, S.-H., 2018. Smart machining process using machine learning: A review and perspective on machining industry. Int. J. of Precis. Eng. and Manuf.-Green Tech. 5 (4), 555–568.

[72] Psarommatis, F., May, G., Dreyfus, P.-A., Kiritsis, D., 2020. Zero defect manufacturing: state-of-the-art review, shortcomings and future directions in research. International journal of production research 58 (1), 1–17.

[73] Kant, G., Sangwan, K.S., 2015. Predictive modelling and optimization of machining parameters to minimize surface roughness using artificial neural network coupled with genetic algorithm. Procedia CIRP 31, 453–458.

[74] Bouacha, K., Terrab, A., 2016. Hard turning behavior improvement using NSGA-II and PSO-NN hybrid model. Int J Adv Manuf Technol 86 (9-12), 3527–3546.

[75] Weichert, D., Link, P., Stoll, A., Rüping, S., Ihlenfeldt, S., Wrobel, S., 2019. A review of machine learning for the optimization of production processes. Int J Adv Manuf Technol 104 (5), 1889–1902.

[76] Yusup, N., Zain, A.M., Hashim, S.Z.M., 2012. Evolutionary techniques in optimizing machining parameters: Review and recent applications (2007–2011). Expert Systems with Applications 39 (10), 9909–9927.

[77] Rinne, J., 2020. Screening datasets for laser welded steel-copper lap joints. Mendeley Data. http://dx.doi.org/10.17632/2s5m3crbkd.2.

[78] Okudan, A., 2018. 3D printer dataset for mechanical engineers. TR/Selcuk University. https://www.kaggle.com/afumetto/3dprinter. Accessed 19 March 2021.

[79] Dittrich, M.-A., Uhlich, F., Denkena, B., 2019. Self-optimizing tool path generation for 5-axis machining processes. CIRP journal of manufacturing science and technology 24, 49–54. https://data.mendeley.com/datasets/smyg6cfwpk/1. Accessed 31 March 2021.

[80] Daimler AG, 2017. Mercedes-Benz greener manufacturing. https://www.kaggle.com/c/mercedes-benz-greener-manufacturing/data. Accessed 19 March 2021.

[81] McCann, M., Li, Y., Maguire, L., Johnston, A., 2008. Causality challenge: Benchmarking relevant signal components for effective monitoring and process control. JMLR: Workshop and Conference Proceedings 6c, 277–288. http://archive.ics.uci.edu/ml/datasets/SECOM. Accessed 19 March 2021.

[82] Brabec, J., Machlica, L., 2018. Bad practices in evaluation methodology relevant to class-imbalanced problems, 4 pp. http://arxiv.org/pdf/1812.01388v1.

[83] Sokolova, M., Japkowicz, N., Szpakowicz, S., 2006. Beyond Accuracy, F-Score and ROC: A family of discriminant measures for performance evaluation. Advances in Artificial Intelligence, 1015–1021.

[84] Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., Bing, G., 2017. Learning from class-imbalanced data: Review of methods and applications. Expert Systems with Applications 73, 220–239.

[85] Harrell, F.E., 2015. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis, Second edition ed. Springer, Cham, Heidelberg, New York, 582 pp.

[86] Bishop, C.M., 1995. Neural networks for pattern recognition. Oxford University Press; Clarendon Press, Oxford, 482 pp.

[87] Provost, F., Fawcett, T., Kohavi, R., 1998. The case against Accuracy estimation for comparing induction algorithms. Proceedings of the 15th international conference on machine learning ICML.

[88] Seliya, N., Khoshgoftaar, T.M., van Hulse, J., 2009. A study on the relationships of classifier performance metrics, in: , 2009 21st IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2009). IEEE, pp. 59–66.

[89] Vishwakarma, G., Sonpal, A., Hachmann, J., 2021. Metrics for benchmarking and uncertainty quantification: Quality, applicability, and best practices for machine learning in chemistry. Trends in Chemistry 3 (2), 146–156.

[90] Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PloS one 10 (3).

[91] Powers, D.M.W., 2007. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. Technical Report SIE-07-001.

[92] Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognition Letters 27 (8), 861–874.

[93] Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143 (1), 29–36.

Biography

Nicolas Jourdan, M. Sc. (*1992) is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt, Germany since 2020. His research interests include the robustness analysis of machine learning models for manufacturing applications.

Lukas Longard, M. Sc. (*1992) is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at at the Technical University of Darmstadt, Germany since 2019. His research interests include the further development and introduction of machine learning to digital shop floor management - a management tool in the production area.

Tobias Biegel, M. Sc. (*1993) is a research assistant and PhD student at the Institute of Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt, Germany since 2019. His research interests include the application of Deep Learning for in-process multivariate statistical process control (MSPC) in manufacturing processes.

Prof. Dr.-Ing. Joachim Metternich (*1968) has been the head of the Institute for Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt, Germany 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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Decisions And Characteristics During The Development Process Of A Software Demonstrator For Data Analysis In Production Logistics

Janine Tatjana Maier[1], Simon Berger[1], Tammo Heuer[2], Peter Nyhuis[2], Matthias Schmidt[1]

[1]*Institute of Product and Process Innovation, Leuphana University of Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany*
[2]*Institute of Production Systems and Logistics, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany*

Abstract

Software demonstrators enable the transfer of theoretical concepts into industrial practice and therefore are used in various research areas. Although the stages of the development of a software demonstrator are already described in literature, it lacks an examination of fundamental decisions and characteristics throughout the development process. Characteristics, such as transparency, hold across all research areas and industries. A specification and extension depends on the individual use case. This paper describes the necessary decisions regarding the framework of a software demonstrator for data analysis in the field of production logistics and the characteristics applicable before and during the actual implementation of the software demonstrator. Key decisions are for example the design of the software architecture or the choice of the programming language. The characteristics equal the requirements to a certain extent and relate directly to the users and to the application itself. Discussions with companies from a wide range of industries revealed underlying conflicts and diverse priorities in terms of the aspects of functionality. Especially the size of the company seems to play an important role regarding the expectations of a software demonstrator. Therefore, the integration of potential future users into the development process of a software demonstrator presents a major advantage. It reduces the number of cycling stages at the end of the development process and increases the willingness to use the software demonstrator after its completion. As the objectives behind software demonstrators can highly differ in the field of production logistics, concepts of demonstrators and software engineering and the related objectives were examined. A demonstrator for determining the order processing strategy serves as an example of how these different approaches interrelate.

Keywords

Software demonstrator; Decision making; Application-related characteristics; User-related characteristics; Order processing strategy

1. Introduction

Demonstrators point out the usefulness of concepts or new technologies. Functioning as testing and experience-generating tools, they help to remove the barriers between theory-based research and actual application in day-to-day industrial practice. Highlighting the potential of applications for a variety of industries, demonstrators often appear in showrooms or at industry fairs. In this case, the appearance as well as the transportability of the demonstrators play a major role. Increasing digitization leads to changes in the nature of demonstrators. There is a noticeable growth in all-software demonstrators and a trend towards combining traditional physical demonstrators with software demonstrators.

Software demonstrators offer the opportunity to automate the calculations of complex models and thus allow access to scientific solutions even for people unfamiliar with the subject. A popular example is a quality control system supported by machine vision [1]. Researchers discuss the use of intelligent systems for planning and controlling production processes for decades. For example, already in 1992 Yang et al. analyzed the similarity of components for the automated generation of process plans [2]. However, the transfer of similar procedures into industrial practice has just started to pick up speed in the last few years. Reasons for this are the increasing availability of data and the rising acceptance of these methods by companies. Nevertheless, a widespread use is still not visible. While companies appreciate the potential of intelligent systems for decision making in production planning and control [3], they also recognize the dependency of the quality of the decision on the quality and integrity of the used data [4]. The analysis of huge amounts of data presents a challenge, especially for small and medium-sized companies. In this area, software demonstrators can provide assistance to a certain extent.

In the field of production logistics various software demonstrators have been developed so far. They range from presenting cause-effect relationships [5] to tools providing assistance to apply machine learning methods [6]. Regardless of the different interpretations and applications of software demonstrators in this field, they are all built upon data analysis. Recent studies indicate that there is still a shortage regarding skilled worker in this field [7]. Companies have the opportunity to compensate this through hiring external experts at high cost. However, the development of software demonstrators for research purposes cannot be outsourced. Instead, this requires a holistic approach including knowledge of theoretical models and scientific research methods. For an adequate support, an examination of the fundamental decisions and characteristics throughout the development process of a software demonstrator is missing. A description of these decisions and characteristics can help to improve the development process and allows focusing on the specifications and extensions for the individual use case.

To incorporate different views on the term software demonstrator and to identify the criteria and influences a literature research based on the established scientific databases Scopus, Web of Science and Google Scholar [8] was carried out. Besides standard literature on software engineering [9, 10], different concepts of demonstrators and the related objective, such as proof of concept, were examined. To evaluate the results and derive conclusions, interviews with companies from a variety of industries were conducted. This included small companies with around ten employees to large multinational companies. Apart from manufacturing companies, logistics service providers, consulting firms and software developers were considered. In the following, a demonstrator for determining the order processing is described to provide an example for the application of the presented general framework. Figure 1 shows the simplified architecture of the software demonstrator for determining the order processing strategy. As an abstract description of the overall system the architecture contains the general functionalities, the modules as well as the methods and patterns. Thus, it acts as a bridge between the requirements and the actual programming [11].

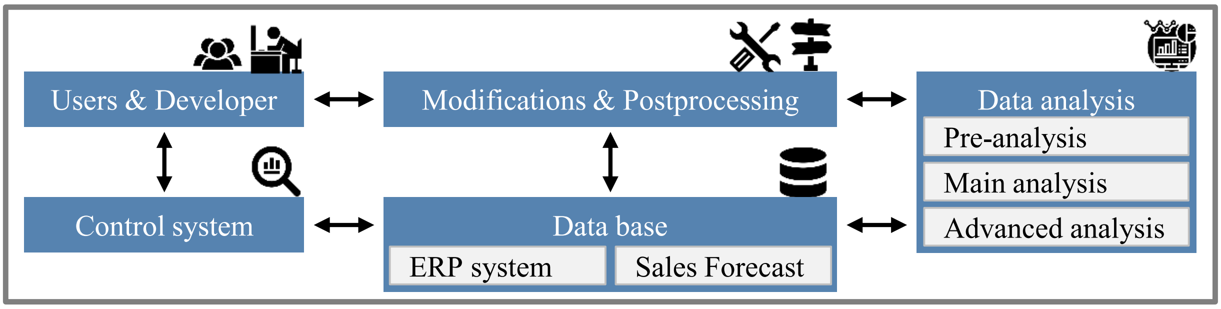


Figure 1: Simplified architecture of a software demonstrator using the determination of the order processing strategy as an example.

The subsequent section describes the development process of a software demonstrator and its challenges. This provides the framework for the analysis of the decisions and characteristics throughout the development process in section three. Furthermore, the characteristics are linked to the application and the users. Lastly, the conclusions of the paper are summarized and future research possibilities are outlined.

2. Development Process of a Software Demonstrator

The principles of a software development process have been examined widely in the literature. Already in 1979, DeMarco described the five phases of software development [12]. Figure 2 shows these phases and highlights the challenges of the application to the development process of a software demonstrator for data analysis in the field of productions logistics. The challenges are derived from literature research, interviews with companies and experience from previous research projects.

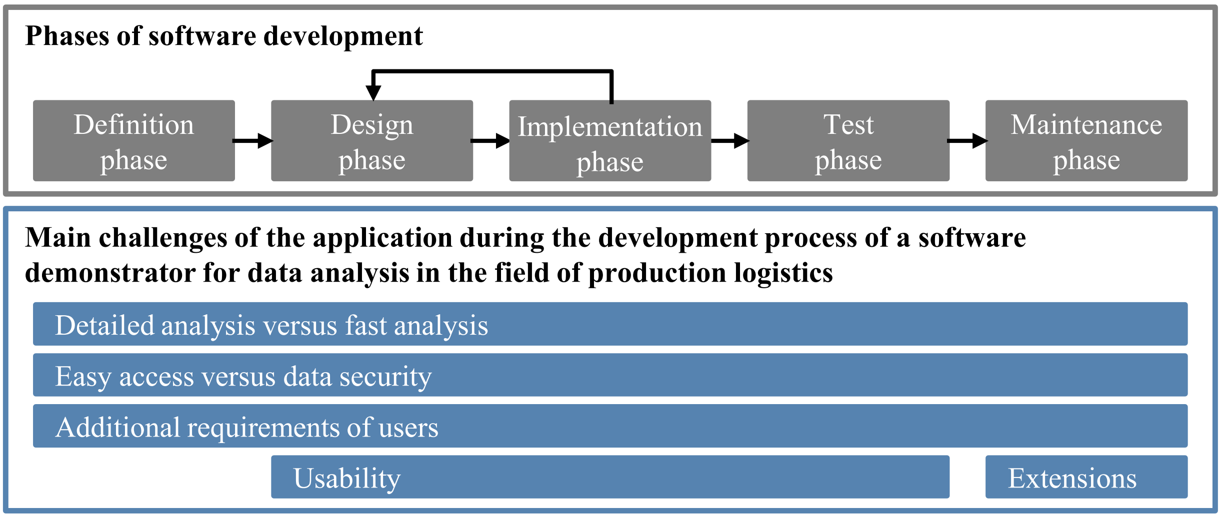


Figure 2: Main challenges for the application of the five phases of software development after DeMarco [12] during the development process of a software demonstrator

The first step is the exact and complete definition of the methodology to be developed. Already during the definition of the functionalities developers face various challenges. An overall objective of software demonstrators is a broad applicability. This leads to a conflict between a detailed analysis and a fast analysis. In addition, easy access as well as high data security standards are required. This is especially important as industrial data are often confidential. These conflicts remain relevant during all phases of the software development process. When transferring theory-based research findings into a software demonstrator, the existing theoretical validated framework can be used as reference during the definition and design phase. As the users of a theoretical model and a software demonstrator can differ additional requirements needs to be taken into account.

Actively involving the interest groups during all phases helps to prevent a deviation between expectations from the actual functionality of the software demonstrator, and thus minimizes the number of cycling stages resulting from non-occurrence. As an attempt to reduce the complexity, researchers further differentiated the definition phase. Balzert provides a good example by distinguishing between requirements in terms of functionality, performance, quality and framework conditions [13]. Models from the area of software development such as the waterfall model [14] or the spiral model [15] are early examples of the importance of a detailed analysis of the realization of the development processes. Such differentiations and descriptions are beneficial. Nevertheless, the identification of the future users is essential as differing needs can have a high impact on the software demonstrator [16].

In this context, it is important to understand the existing types of interest groups and evaluate the benefits of extending the system. Existing software demonstrators, such as the FRDISI [17] or the 4SECURail demonstrator [18], point out diverse interest groups for their respective areas and mainly base the general architecture of their demonstrators on these. This supports the assumption that a key factor behind the generation of the so-called usability and the associated user acceptance is the involvement of future users in the entire development process. The success of the incorporation of the results provided by the software demonstrator in a company highly depends on the acceptance of the employees regarding new technologies or processes. A human-centered design process helps to incorporate the user in an appropriate way. Besides other characteristics, a human-centered design process understands and actively involves the users and according to ISO 13407 [19]. Combining approaches of human-human and human-computer interaction research creates the basis for identifying the actual underlying criteria of such a process [20]. Trending topics such as artificial intelligence push this idea further [21]. Depending on the individual use case, additional approaches from fields like decision support systems [22] or simulation [23] can complement these criteria.

The design of user interfaces directly relates to usability. In 1993, Nielsen defined practices for usability engineering [24]. For novel technologies and concepts, trustworthiness is one of the most critical aspects. Mathis et al. show how personalized interfaces create trustworthiness, persuasion and even affection towards automated transport vehicles [25]. Similarly, a study on mobile services adoption points out the importance of trust and the user's desire to be in control [26]. With increasing digitalization and the growing interest in topics like artificial intelligence, the requirements of users broaden and result in design principles for specific topics, such as Industry 4.0 components [27]. The rising level of expectations has a strong impact on the data requirements. Data should not only be available, but also transparent and able to process in real-time [28]. The usability of data also demands easy data inputs and modifications of the data base [26]. For example, well-known technologies such as radio frequency identification provide opportunities for real-time data processing and therefore enable dynamic decision making [29].

Regardless of the degree of involvement of future users, an evaluation of the usability before, during and after the implementation of a system is mandatory. Already in 1997, Lin et al. proposed an index to measure the usability of software interfaces [30]. They take a wide range of requirements into account such as the fit to the user's expectations, the consistency, the flexibility, the learnability and the provided guidance. A variety of methods exists to conduct evaluations throughout the development process. As a way to eliminate obvious usability problems prior to an actual test, it is possible to carry out a cognitive walkthrough [31]. In terms of an actual test, Holzinger defines thinking out loud as one of the most valuable methods to identify shortcomings regarding usability [32]. Scholtz provides a good overview of evaluation methods [33].

3. Decisions and Characteristics

Throughout the development process of a software demonstrator, multiple decisions are required. The decisions are mainly driven by the defined objectives, but also directly interrelate with the users and the application itself. In the following, possible influences and criteria are outlined. Figure 3 visualizes the key decisions and characteristics during the development process of a software demonstrator for data analysis in the field of production logistics. It shows the underlying conflicts and highlights the importance of a continuous analysis, validation, and verification. The characteristics partly equal the requirements. Characteristics can influence the development process, but their existence is not necessary in every case. Requirements are considered mandatory for the development process of a software demonstrator. For example, the project lead-time can be seen as an externally given specification and have a strong impact on the entire developing process. Therefore, it classifies as a characteristic and as a requirement. The pre-existing infrastructure influences the development process of a software demonstrator, but the existence of an instructor prior to the development process is not necessary. Thus, it is only a characteristic.

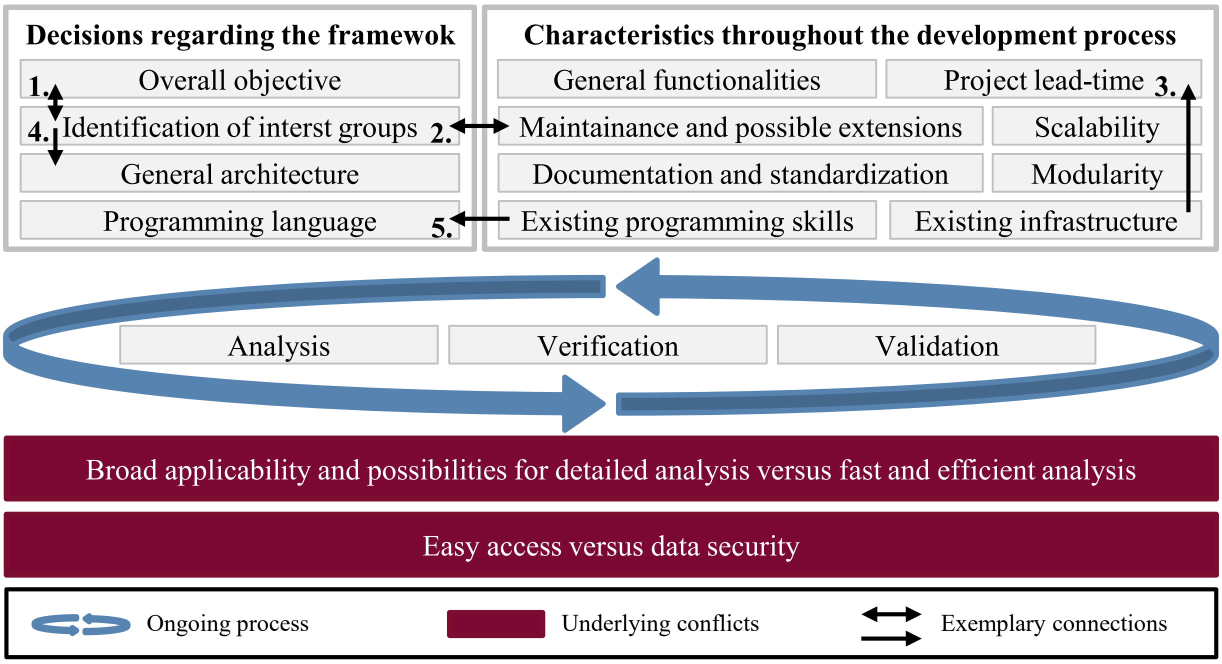


Figure 3: Decisions and characteristics, ongoing processes and underlying conflicts during the development process of a software demonstrator

In the following, the individual aspects are described in more detail and examples of connections based on a software demonstrator for determining the order processing strategy are given. It should be noted, that all decisions and characteristics are directly or indirectly connected. Defining the overall objective of a software demonstrator represents the beginning of the development process. The development requires a considerable amount of resources such as time, effort and money, and is a never-ending process in case the software demonstrator changes along with evolving companies over the years. To limit the effort the project lead-time and the general functionalities should be specified prior to any decisions. In contrast to commercial software development projects, there is no contract detailing these aspects. Nevertheless, building a software demonstrator based on an existing theoretical model has a major advantage. The concept has usually already been validated, for example by simulation studies. Thus, the risk of a project failure is comparatively low. The overall objective is directly related to the general functionalities and the project lead-time. The definition of the overall objective leads to the obvious interest group, the users. Identifying additional interest groups and incorporating their requirements can lead to adjustments of the objective definition (figure 1, arrow 1). The objectives of all interest groups should withstand in the real world. In the case of programming software, a wide range of possible requirements exists and thus additional support is required to merge the requirements together. Nuseibeh and Easterbrook suggest a roadmap to specify the requirements of software products [34]. Moreover, further classifications of the requirements such as a distinction between functional and non-functional can be beneficial. Aurum and Wohlin provide a list with examples for different requirement classifications [35]. Risk management approaches suggest to confront the interest groups in case of non-occurrence and to rank all requirements. Disguising non-occurrence will simply postpone the problem and even intensify it [36]. Therefore, the first and probably most continuously present decision is the selection of the interest groups. This calls for an additional distinction between the users and the generally possible interest groups. For the majority of cases the interest groups equal the users. Nevertheless, people do not need to be direct users to influence the development process. This includes people related to the data collection and preprocessing as well as people whose work is influenced by the results of the demonstrator. Neglecting to align the demonstrator with data acquisition and preprocessing results in the need of additional data standardization, cleansing and format transformation. Although this is possible without programming skills, for example through software such as KNIME [37], these additional process steps affect the usability of the demonstrator negatively.

Apart from that, it is important to know about the data already collected but so far not used and the cost-benefit ratio between the gathering of additional data and the quality of the results from the demonstrator. Involving the responsible parties eases the estimation of the necessary time and effort due to existing experience and even helps to reduce barriers regarding possible future extensions of the data collection. The interpretation of data requires a considerable amount of effort as well. Therefore, the results from the demonstrator need to be easy to evaluate and appropriate actions should be derivable without a high amount of effort and the consultation of experts [22]. A way to achieve this is the identification of tools, which comely exist in companies and possibly will be used for postprocessing the data of the demonstrator. This also raises the question who will maintain and potentially upgrade the system from a long-term perspective. Unless these tasks belong to the developer himself, more people need to be actively included in the development process (figure 1, arrow 2). Especially important in this context are a high level of detail in the documentation and the use of standards. On top of simplifying maintenance and extensions, standards provide the foundation for scalability [38]. Individual use cases imply different levels of detail and complexity. The main challenge is to enable a broad applicability and possibilities for detailed analysis, but at the same time provide a fast and efficient analysis [22].

The demonstrator should be applicable to both small and very large companies [23]. The number of people potentially connected to the demonstrator as well as the expectations tend to increase along with the company size. Small companies tend to focus on aspects such as reusability due to the usually more limited resources available for development. Using pre-defined elements can help to achieve a higher level of detail while meeting the defined project lead-time (figure 1, arrow 3). In addition, not needing to start modeling from scratch reduces the occurrence of errors [22]. On the other hand, fitting a new system to an existing infrastructure can result in a less efficient analysis. Aiming for a reusable and scalable system generates interest for more companies and therefore is a common objective for demonstrator developments [38]. Adaptability to different companies includes different industries as well as a variety of products and processes. Thus, modularity is one of the key aspects to achieve a scalable and reusable system [23]. In addition, a modular architecture is beneficial in case users want or are expected to make individual adjustments to the demonstrator. The challenge is to balance a universally applicable solution, a high level of detail and easy to follow calculation. Unless the demonstrator is considered as a sort of black box, a common ground regarding terms and data flow structure is essential [39]. After determining the interest groups, the general architecture is designed with regards to the identified requirements (figure 1, arrow 4).

The next step is the selection of the programming language. Consideration the programming skills of the developer, the users and if possible the long-term system manager helps to reduce the project lead-time as well as the maintenance and modification effort. On the downside, it heavily restricts the programming language (figure 1, arrow 5). As a result, a programming language or evaluation method that is actually more suitable for the application may be rejected due to a lack of experience. For example, R is a common software for statistical data analysis, but requires a considerable amount of knowledge [40]. In addition, the question of how to balance easy access and data security arises. Web-based solutions ease the linkage to other systems and enable a wide user base. Nevertheless, in the case of processing confidential data, companies commonly favor non web-based solutions. Data and users are also key aspects in the verification and validation. Only a high level of data quality and the correct usage lead to the desired result. A verification is especially important in the case of modular systems as it ensures the correctness of both the individual modules and the overall system. Just like models, demonstrators represent abstractions of real-world problems or systems. They support decision making by helping to estimate the outcome prior to an actual implementing [41]. Therefore, the validation by a real industrial application is mandatory to control the level of abstraction [22].

In case of the software demonstrator for determining the order processing strategy the central challenge was to connect the various requirements resulting from the fact that the order processing strategy interacts with upstream strategic aspects and at the same time with downstream production planning and control tasks. As a result, not only different companies and their diverse production processes but also numerous divisions within this companies had to be involved in the development process. One way of achieving this was to set up a working group for the companies to exchange views and experience amongst each other. In addition, several companies were visited to generate a common understanding of their production processes and the associated procedures.

4. Conclusions and Outlook

Transferring theory-based research findings into the day-to-day industrial practice is challenging. Software demonstrators can help to speed up this process by linking theory-based research to the real world problems of companies. Nevertheless, they are only temporary solutions. Companies might use a software demonstrator as a starting point for their own customized developments. Researchers can upgrade the software demonstrator as they gain new insights or connect it to other demonstrators to take a step towards a smart factory. Existing literature often neglects the decisions and characteristics during the development of a software demonstrator or focuses on a very specific case of implementation. Interviewing companies from a variety of industries revealed a rising level of expectations for software demonstrators in the field of productions logistics aligned with the company size. Large companies tend to have more people involved in the development process and more pre-existing infrastructure. In this case, the decision, which requirements should be prioritized, and which requirements should not be taken into account, is essential. Employees in small and medium-sized companies usually have less time available to take an active part in the development process. This shifts the focus to deciding when users are actually needed and to what extent testing with fictitious data or freely available data from other areas is sufficient. Incorporating different sized companies from various industries in the development process ensures a broad applicability. Thus, a software demonstrator should be designed to fit shifting objectives. Even though the decisions and characteristics are influenced by the company size, it does not change their strong interrelations and the underlying conflicts. Only a continuous analysis, verification and validation alongside a clear position between the conflicting objectives can lead to the development of a suitable software demonstrator. As the development process of a software demonstrator differs from a classical software engineering process, further research is required. Analyzing the reusability of pre-defined software elements could help to standardize the development process. For example, examining the wide range of existing projects containing software demonstrators could help to highlight similarities and at the same time specify the decisions and characteristics tendencies for different research areas.

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 Energy (BMWi) 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] Louw, L., Droomer M., 2019. Development of a low cost machine vision based quality control system for a learning factory. Procedia Manufacturing 31, 264-269. https://doi.org/10.1016/j.promfg.2019.03.042

[2] Yang, H., Lu, W.F., Lin, A. C., 1992 Intelligent Process Planning Using a Machine Learning Approach. IFAC Proceedings Volumes 25 (28), 147-151. https://doi.org/10.1016/S1474-6670(17)49482-1

[3] Schmidt, M., Maier, J.T., Grothkopp, M., 2020. Eine bibliometrische Analyse - Produktionsplanung und -steuerung und maschinelles Lernen (english title: PPC and machine learning - A bibliometric analysis). Wt werkstattstechnik online 110 (4), 220-225.

[4] 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 66 (1), 425-428. https://doi.org/10.1016/j.cirp.2017.04.003

[5] Härtel, L., Nyhuis, P., 2018. Systematic Data Analysis in Production Controlling Systems to Increase Logistics Performance, Schmitt R., Schuh G. (eds) Advances in Production Research. Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP), Aachen, November 19-20, 2018. https://doi.org/10.1007/978-3-030-03451-1_1

[6] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Google Brain, 2016. TensorFlow: A System for Large-Scale Machine Learning. In 12th USENIX symposium on operating systems design and implementation (OSDI 16), pp. 265-283

[7] Rammer, C., Bertschek, I., Schuck, B., Demary, V., Goecke, H., 2020. Einsatz von Künstlicher Intelligenz in der Deutschen Wirtschaft: Stand der KI-Nutzung im Jahr 2019. ZEW-Gutachten und Forschungsberichte.

[8] Martín-Martín, A., Orduna-Malea, E., Thelwall, M., López-Cózar, 2018. Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. Journal of Informetrics 12 (4), 1160-1177. https://doi.org/10.1016/J.JOI.2018.09.00

[9] Bourque, P., Dupuis, R., Abran, A., Moore, J.W., Tripp, L., Wolff, S. 2002. Fundamental principles of software engineering – a journey. The Journal of Systems and Software 62 (1), 59–70

[10] Juristo, N., Moreno, A.M., 2013. Basics of software engineering experimentation. Springer Science & Business Media.

[11] Garlan, D., 2000. Software architecture: a roadmap, in: Proceedings of the Conference on the Future of Software Engineering, pp. 91-101.

[12] DeMarco, T., 1979. Concise Notes on Software Engineering, Yourdon Press, New York.

[13] Balzert, H., 2009. Lehrbuch der Softwareentwicklung. Basiskonzepte und Requirements Engineering, 3rd ed. Spektrum, Heidelberg.

[14] Adenowo, A.A., Adenowo, B.A., 2013. Software Engineering Methodologies: A Review of the Waterfall Model and Object-Oriented Approach. International Journal of Scientific & Engineering Research 4 (7), 427-434.

[15] Boehm, B.W., 1988. A spiral model of software development and enhancement. Computer 21 (5), 61-72. doi: 10.1109/2.59.

[16] Boehm, B., Bose, P., Horowitz, E., Lee, M.J., 1999. Software Requirements Negotiation and Renegotiation Aids: A Theory-W Based Spiral Approach. Proceedings - International Conference on Software Engineering. 10.1145/225014.225037.

[17] Coulibaly, M., Belkhala, S., Errami, A., Medromi, H., Saad, A., Rouissiya, M., Jaafari, A., 2018. Development of a demonstrator «smart-parking», in: 2018 19th IEEE Mediterranean Electrotechnical Conference (MELECON), IEEE, pp. 172-176. doi: 10.1109/MELCON.2018.8379088.

[18] Basile, D., ter Beek, M.H., Fantechi, A., Ferrari, A., Gnesi, S., Masullo, L., Mazzanti, F., Piattiona, A., Trentini, D., 2020. Designing a demonstrator of formal methods for railways infrastructure managers, in: Tiziana, M., Bernhard, S. (Eds.), International Symposium on Leveraging Applications of Formal Methods, Springer, Cham, pp. 467-485.

[19] ISO 9241-210, 2019. Human-centred design processes for interactive systems. International standard, 2nd ed. International Standardization Organization

[20] den Os, E., Boves, L., 2003. Towards ambient intelligence: Multimodal computers that understand our intentions, in: Proceedings of eChallenges e-2003, pp. 22-24.

[21] Mou, Y., Xu, K., 2017. The media inequality: Comparing the initial human-human and human-AI social interactions. Computers in Human Behavior 72, 432-440. https://doi.org/10.1016/j.chb.2017.02.067

[22] Blume, S.A., 2020. Resource Efficiency in Manufacturing Value Chains, Springer, Cham.

[23] Randell, L.G., Bolmsjo, G.S., 2001. Database driven factory simulation: a proof-of-concept demonstrator, in: Proceeding of the 2001 Winter Simulation Conference (Cat. No. 01CH37304), vol. 2, IEEE, pp. 977-983.

[24] Nielsen, J., 1993. Usability Engineering. Academic Press, Inc. ISBN 0-12-518405-0

[25] Mathis, L.A., Diederichs, F., Widlroither, H., Ruscio, D., Napoletano, L., Zofka, M. R., Viehl, A., Fröhlich, P., Friedrich, J., Lindström, A., Thorslund, B., Litke, A., Papadakis, N., Bakalos, N., Hernandez-Jayo, U., Espié, S., cavallo, V., Panou, M., Gaintnidou, E., Bekiaris, E., (2020). Creating informed public acceptance by a user-centered human-machine interface for all automated transport modes, in: Rethinking transport. 8th Transport Research Arena TRA 2020, April 27-30, 2020, Helsingfors, Finland.

[26] Kaasinen, E., 2005. User acceptance of mobile services: Value, ease of use, trust and ease of adoption.

[27] Hermann, M., Pentek, T., Otto, B., 2016. Design principles for industrie 4.0 scenarios, in: 2016 49th Hawaii international conference on system sciences (HICSS), IEEE, pp. 3928-3937.

[28] Mantravadi, S., Jansson, A. D., Møller, C., 2020. User-Friendly MES Interfaces: Recommendations for an AI-Based Chatbot Assistance in Industry 4.0 Shop Floors, in: Asian Conference on Intelligent Information and Database Systems, Springer, Cham, pp. 189-201.

[29] Zhong, R.Y., Li, Z., Pang, L.Y., Pan, Y., Qu, T., Huang, G.Q., 2013. RFID-enabled real-time advanced planning and scheduling shell for production decision making. International Journal of Computer Integrated Manufacturing 26 (7), 649-662. https://doi.org/10.1080/0951192X.2012.749532

[30] Lin, H.X., Choong, Y., Salvendy, G., 1997. A proposed index of usability: A method for comparing the relative usability of different software systems. Behaviour & Information Technology 16 (4-5), 267-277. https://doi.org/10.1080/014492997119833

[31] Jeffries, R., Miller, J.R., Wharton, C., Uyeba, K.M., 1991. User interface evaluation in the real world: a comparison of four techniques, in: Procedings of ACM CHI'91 conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, pp. 119- 124.

[32] Holzinger, A., 2005. Usability engineering methods for software developers. Communication of the ACM 48 (1), 71-74. https://doi.org/10.1145/1039539.1039541

[33] Scholtz, J., 2004. Usability evaluation. National Institute of Standards and Technology

[34] Nusbeih, B., Easterbrook, S., 2000. Requirements Engineering: a Roadmap, in: Proceedings of the conference on The future of Software engineering - ICSE '00, vol. 21, no. 5, pp. 61-72. doi: 10.1145/336512.336523

[35] Aurum, A., Wohlin, C., 2005. Engineering and Managing Software Requirements, Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-28244-0

[36] DeMarco, T., Lister, T., 2003. Risk management during requirements. IEEE software 20(5), 99-101

[37] Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Thiel, K., Wiswedel, B. (2009). KNIME-the Konstanz information miner: version 2.0 and beyond. AcM SIGKDD explorations Newsletter 11(1), 26-31. https://doi.org/10.1145/1656274.1656280

[38] Alexandersson, J., Becker, T., Engel, R., Löckelt, M., Pecourt, E., Poller, P., Reithinger, N., 2004. Ends-based dialogue processing, in: Proceedings of the 2nd International Workshop on Scalable Natural Language Understanding (ScaNaLU 2004) at HLT-NAACL 2004, pp. 25-32.

[39] Maybury, M.T., Wahlster, W., 1988. Readings in Intelligent User Interfaces. Morgan Kaufmann, San Francisco.

[40] Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. Journal of computational and graphical statistics 5(3), 299-314.

[41] Law, A.M., Kelton, W.D., 1991. Simulation modeling and analysis, McGraw-Hill, New York.

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.

Simon Berger (*1992) studied industrial engineering at the Hochschule Ravensburg-Weingarten University of Applied Science and the Leuphana University of Lüneburg. Since 2021 he works as Production Process and Technology Engineer in the company Franatech GmbH.

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.

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.

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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

A two-stage Tabu Search for multi-objective facility layout problem

Paul Aurich[1], Marc Speckmann[1], Christian Böning[1], Malte Stonis[1]
[1]*IPH – Institut für Integrierte Produktion Hannover gGmbH, Hannover, Germany*

Abstract

In this paper, a new solution for the facility layout problem is presented. The approach was integrated into a planning software. The aim of the MeFaP research project was mainly the development of a user-friendly decision support regarding the facility layout problem for small and medium-sized companies. Therefore, a realistic modelling of the planning problem was focused on. Thus, a path planning with area allocation was integrated, for example. The metaheuristic Tabu Search was selected as a solution approach. To ensure an efficient optimisation, the optimisation is performed in two steps, once without and then with route planning. The experiments were performed with the objectives material flow distance, temperature and cleanliness, which are briefly described. The results of the experiments were compared with current solution approaches.

Keywords

multi-objective facility layout problem; factory planning; multi-criteria optimization; metaheuristic; tabu search; planning software

1. Introduction

In the research project MeFaP, an optimization approach for quantitative, multi-criteria facility layout planning was developed [1–3]. More precisely for the facility layout problem (FLP), which is defined as the arrangement of facilities in a factory site, to achieve the best possible fulfilment of objectives while taking restrictive constraints into account. The project aimed to provide an user-friendly decision support for non-professional users. Furthermore, quantitative layout evaluation formulas, which were developed in the preliminary project QuaMFaB [4], were examined regarding their applicability for a multi-objective facility layout problem. Since currently used optimization approaches are mostly limited regarding their objective function. Often only one objective is optimized, the material flow distance respectively material handling costs (cf. [5]). A comprehensive survey of current approaches is given by DRIRA ET AL., HOSSINI-NASAB ET AL. as well as SHARMA AND SINGHAL [6,5,7]. In facility layout planning, however, other objectives respectively objective fields, such as changeability, communication flow or occupational health and safety standards (like temperature and cleanliness), are in many cases also relevant for planning (cf. [9,8]). For this reason, a multi-criteria solution approach was developed to enable holistic optimization of facility layouts. Due to the structure of the evaluation formulas, a discrete layout presentation was necessary (cf. [4]). This means that the factory floor is covered by a grid of square cells with equal size, which is variable and can be adapted by the user, to model specific planning projects. The facilities are also covered with the grid. In this way, different information can be assigned to individual cells of the factory floor or facilities. For example, the availability of a medium such as electricity or the lighting intensity of daylight can be assigned to a cell of the factory floor. A corresponding demand can be assigned to the cells of a facility. By positioning a facility on the grid, the superimposed cells are linked together. The layout is evaluated by comparing the demand of the facility cell and the supply of the factory floor cell. Due to the discrete representation, the

FLP is modelled as a quadratic set covering problem (QSP). The facilities are characterized by a regular shape and a fixed aspect ratio. In addition, the facilities have fixed pick-up and drop-off points as well as spatial orientation. Overlapping of facilities is not allowed. The optimization problem is also characterized as an open field layout. During the optimization, the position of the facilities and the necessary path structure for the layout is arranged.
According to prevailing opinion in the literature, a continuous problem formulation is better suited to find an optimal solution for the FLP than a discrete problem formulation [6,10]. Due to the grid-based problem formulation, facilities could not be represented with their exact size, and have to be adapted to the cell size of the grid. This usually leads to an (slight) enlargement of the facilities. Because of the insufficient detailing, it might not be possible to consider exact pick-up and drop-off points [6]. Thereby a path planning is also shown as not possible. As the main disadvantage of a discrete problem formulation, it is argued that the "real" optimal respectively continuous position of the facilities cannot be found because of the limited solution space [12,11].
This argumentation ignores the fact that the level of detail depends on the definition of the cell sizes. As smaller the cells are, as greater the detailing is. Furthermore, it is neglected, that the often-used fixed size and shape of facilities is based on the assumption, that this information is known a priori. In the research project, all assumptions were verified with manufacturing companies. According to the predominant opinion of the companies, a fixed representation of facilities is sufficient in the phase of block layout planning. Because the final design of the facilities is done in the subsequent realization planning. Accordingly, an approximate positioning of pick-up and drop-off points is also sufficient, since these points are also finalized in a later planning step. BOCK AND HOBERG provide a realistic layout planning with a discrete approach, where paths are considered [13].
Within the project, a decision was made against a continuous problem formulation. In existing continuous approaches, pick-up and drop-off points as well as paths are taken into account [5,11]. However, the paths usually pass along the outer edges of the facilities or a rectangular path routing is assumed, whereby paths in some cases even cross trough other facilities [14,12]. The necessary path area is often not considered in the layout [16,15]. Consequently, an optimal positioning of the facilities is achieved, but this is not realizable without considering the path area. Therefore, current approaches are often not suitable for realistic layout planning. KLAUSNITZER ET AL. provide a continuous approach, in which path areas are planned [14].
In the project, a decision support system should be developed that suggests companies the most optimal, but also realistic layout variants. As mentioned before, the discrete layout representation was necessary to integrate the previously developed evaluation formulas. This can be realized with the chosen approach.

2. Optimization Approach

This paper presents an optimization approach based on the metaheuristic tabu search. In the following, the implementation of tabu search, as well as the associated procedures for neighbourhood search, are presented. Furthermore, the objectives are explained. First, the problem formulation is explained in more detail. In the description of the optimisation approach, the developed software is also referenced. For example, user input will be described. Figure 1 shows the basic process of optimization.

2.1 Implementation of the problem representation

For discrete problem formulation, the factory surface is covered by a grid of square cells of equal size. The size of the cells can be defined by the user in integer steps. The shape of the factory floor can be modelled by hiding cells. In this way, non-regular shapes can also be represented.
Restrictions can be assigned to each cell of the factory floor. These include ceiling height, floor load, ceiling load. Additionally, information regarding media can be assigned to the cells. Media do not represent restrictions during the optimization but can be used as an objective. The number of media can be defined by

the user (e.g. the availability of water, electricity, or compressed air). Furthermore, restrictive areas can be defined for different types of facilities or departments (e.g. production, assembly, warehouse). If one or more restricted areas have been defined for a facility type, it is only possible to position the associated facilities in this area. If no restrictive area is defined, facilities can be positioned anywhere, even in restrictive areas of other facility types. Additionally, it is possible to position fixed path cells.

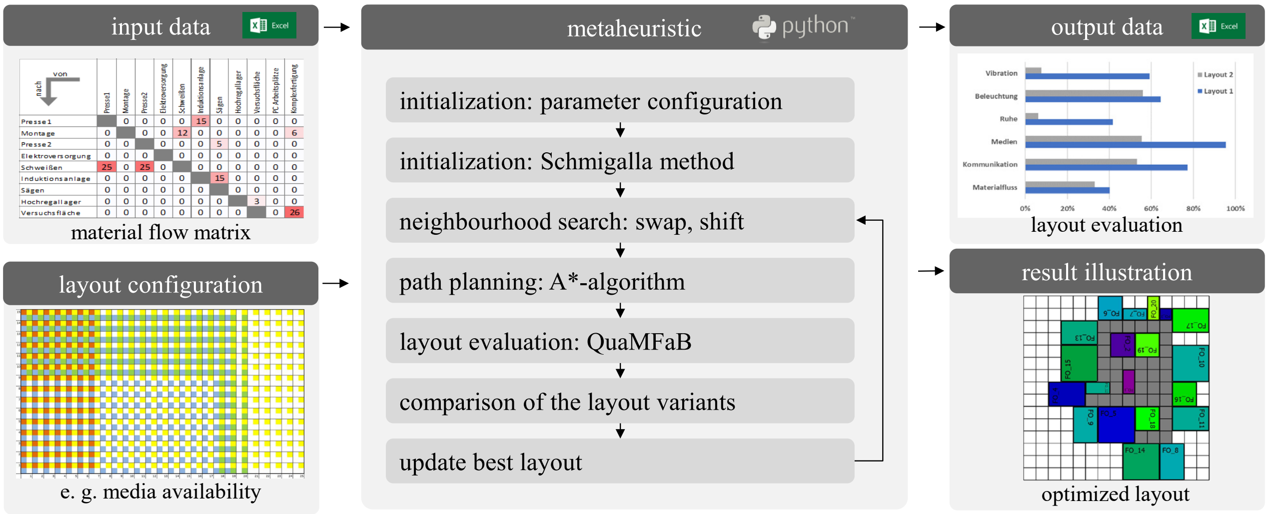


Figure 1: abstract optimisation process

The position of the facilities in the layout is defined by their upper left corner. Facilities can be rotated in four steps (0°, 90°, 180°, 270°). Furthermore, it is possible to mirror facilities on the vertical axis. Rotation and mirroring affect the position of pick-up and drop-off points in the layout. The definition of pick-up and drop-off points is also done by the user within the setup. When positioning is not performed, pick-up and drop-off points are placed in the upper left corner at a rotation of 0°. The position of facilities in the factory floor can be fixed by the user. For example, machines with special fundaments could be considered in planning without changing their position. The fixing was a requirement of the consulting companies.

2.2 Metaheuristic

Tabu Search (TS) is a local search technique developed by GLOVER [17]. The method is based on the tabu setting of transformation steps. A transformation step is the repositioning of a facility in the layout. With an iteration of the tabu search, a repositioning is done for each facility. This means that as many new layouts are created as there are facilities. In an iteration, the same neighbourhood search method is performed for all facilities. This is a partial search in the neighbourhood. The best transformation step is accepted as new solution, set as tabu, and added to a tabu list. If a transformation step subsequently leads to the best solution in the neighbourhood again, it is not selected, but the next best neighbour is selected. The length of the tabu list (TL) controls how long a solution remains tabu. When the tabu list is full, the oldest entry is deleted. The tabu list avoids circling around a local optimum. If a tabu neighbour is better than the best-known solution, the tabu mechanism is bypassed; this is a so-called aspiration criterion.

2.3 Initialization

Two initialization methods were implemented. An adaptation of the Schmigalla method and a random based positioning. During the initialization, only feasible layouts are generated. Accordingly, all restrictions are strictly respected. If it is not possible to position a facility during initialization, the layout is deleted, and the method is executed again. The user can select the initialization method. Various experiments have shown that the random-based method is more likely to generate a valid initial layout when a high degree of space utilization is given.

In the adaptation of the **Schmigalla method**, first, the order of the facilities is determined based on their material flow relations. Then the first facility is positioned in the middle of the factory floor. Subsequently, it will be tried to arrange the next facility to the left, right, above or below the previously positioned facility. The direction is selected randomly. This process is repeated until all facilities are positioned.

At the beginning of the **random-based initialisation**, the facilities are sorted in descending order according to their surface area. Based on this sequence, the facilities are placed at a random position. Rotation and mirror are also randomly selected. Since the largest facilities are positioned first, it is reliable to find a feasible solution even with a high space utilization ratio.

2.4 Neighbourhood search methods

Two neighbourhood search (NHS) methods were implemented. The selection of which neighbourhood search method is used in the optimization can be controlled by the user. For this purpose, a distribution can be set, which must sum up to 100%. During the optimization, a random number between 0 and 1 is chosen, the NHS is selected by comparing it with the distribution.

The first method is an adaptation of the **Local Reallocation Search** (LRS) by BOCK UND HOBERG [13], which tries to reposition a randomly selected facility at a random position nearby. The rotation and mirroring of the facility are also determined randomly. If a check confirms that all restrictions are met, the facility is positioned, and the method is left. If restrictions are violated, a new random position, rotation and mirroring are chosen and checked again. This procedure repeats until an a priori set iteration maximum is reached ($LRS_{Iteration}$). By default, this is set to 100 iterations, however, it can be changed by the user. If the iteration maximum is reached, a new facility is randomly selected, and the process starts again. The maximum permissible distance to the original position is used as a control parameter of the LRS ($LRS_{StepSize}$). By default, the value is set to 10 cells and can also be changed by the user. Good results are achieved if the value reflects approximately the edge length of a medium-sized facility related to the data set.

In the **Open Area Search** (OAS), a facility is randomly selected first. Then the size of all open areas in the layout is calculated in which the selected facility could fit in. The selection of the open area can be performed in two different ways. In the first variant, the determined open areas are sorted in descending order of size. This is an attempt to prevent large open areas from being occupied by small facilities. In the second variant, the open areas are sorted in ascending order concerning the distance to the current position of the facility. After sorting, the method checks if the facility could be positioned with respect to the restrictions. If this is the case, the facility is repositioned, and the method is left. If restrictions would be violated, the next open area from the list is checked. If no suitable open area is available, a new facility is randomly selected, and the process starts again. Figure 2 shows the application of the OAS concerning facility $f = 3$, with the size-oriented sorting variant. In this example, open area$_3$ would be selected first for the check.

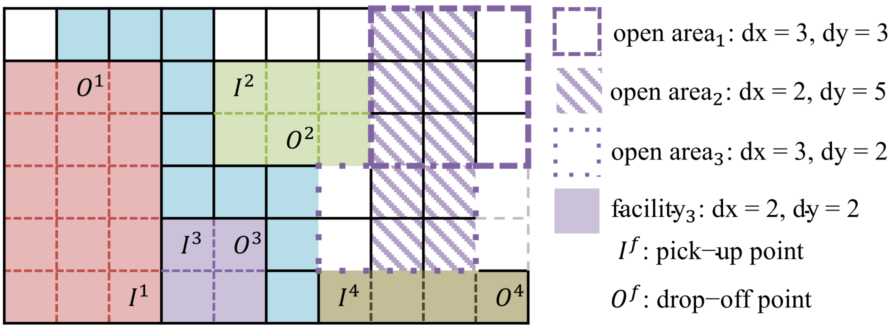


Figure 2: Example of possible open areas for facility $f = 3$

2.5 Path planning

Two methods were used to calculate the transport path length, the A* algorithm and the Euclidean distance. The **A* algorithm** is a well-known and frequently used shortest path algorithm first published by HART ET

AL. [18]. Compared to other shortest path algorithms, it requires less computing power and memory. The A* algorithm always finds the shortest path between two points. In the present case, this concerns the pick-up and drop-off points of the facilities between which a material flow is necessary. Free cells, fixed path cells and local path cells are available for path planning. Local path cells are cells that have already been used by a path between other facilities. These are accordingly planned in the layout. The use of various cell types has different "costs". The default values of the route costs are: fixed route cells = 1, local route cells = 1.5 and free cells = 3. Accordingly, the A* algorithm would take a diversion of 3 cells to use fixed route cells instead of free cells. It always determines the shortest respectively the cheapest path. In this way, it is avoided that multiple parallel paths are created in the layout. Path costs can also be modified by a user.

The **Euclidean distance** is the direct connection between two points. If the Euclidean distance is used to calculate the transport path lengths, local path cells are not planned in the layout. The calculation of the method is many times faster than the A* algorithm.

2.6 Two-stage optimization

In order to reduce the computing times, a two-stage optimisation was implemented. The functionality of the applied heuristics is not affected. The acceleration is mainly based on the selection of the path planning approach. In the first optimisation stage, the Euclidean distance is used to calculate the transport path lengths. In the second stage, the A* algorithm is used. In this way, about 10 to 15 times more layout variants can be evaluated at the same time. The control parameters of the optimisation stages can again be set by the user. These should be selected in such a way that the first stage takes up a much larger share of the total computing time. During the experiments, a ratio of 20:1 was chosen.
Between the stages, the layout must be re-initialised to enable path planning with the A* algorithm. This is because, in the first stage, the pick-up and drop-off positions of the facilities can be blocked by other facilities or the external walls of the factory. This is not possible when using the A* algorithm, because otherwise no path can be found. The intermediate initialisation of the layout is based on the Local Reallocation Search, trying to place the facilities at the same position as before. In contrast to the behaviour with the Neighbourhood Search, the positioning is not random, instead, it moves outwards step by step from the original position. Rotation and mirroring are also not randomly but systematically varied.

2.7 Layout evaluation

The evaluation of the layout variants is based on the evaluation formulas designed in the preliminary project [4]. These were adapted or improved in the current project [1]. Scaling of the objectives is not necessary. The result values of all evaluation formulas are percentage values with a value range of $\{0, 100\}$. This results from the comparison of layout variants. An intermediate evaluation result of a current layout variant is compared with the best-known intermediate evaluation result of the respective objective. The best-known evaluation result is constantly updated during the optimisation process if there is a better solution. This modelling ensures direct comparability or transferability. The modelling also allows the combination of objectives that have to be minimised and maximised. However, a weighting factor ω is applied to the combination. It can also be configured by the user. In the following, only the evaluation formulas relevant for the paper are presented.

f: facility
m: material flow between a pair of facilities
v: layout variant

Material flow distance

$$MF_v = \left(\frac{MFd_{min}}{MFd_v}\right) \cdot 100\% \quad (1)$$

MF_v: objective function of the material flow distance for variant v
MFd_v: material flow distance of variant v in metres
MFd_{min}: minimum material flow distance of all variants

$$MFd_v = \sum_{m=1}^{M}(MFI_m \cdot d_m) \tag{2}$$

d_{m}: distance of material flow m in metres
MFI_m: transport intensity of material flow m

Temperature

The objective temperature is an objective function which must be maximised. Thus, influencing facilities should be positioned separately from each other. For this purpose, a temperature factor TE_f is assigned to each facility. The factor is freely selectable but must follow a predefined scheme: positive values represent temperature emission, zero means neutrality and negative values indicate temperature sensitivity. The more the values of two facilities differ, the greater is the distance requirement (4). Dependencies must be indicated with a sign change.

$$Temp_v = \left(\frac{Tempd_v}{Tempd_{max}}\right) \cdot 100\% \tag{3}$$

$Temp_v$: objective function of the temperature-sensitivity in variant v
$Tempd_v$: temperature-sensitive distance of variant v
$Tempd_{max}$: variant v with the maximum temperature-sensitive distance

$$Tempd_v = \frac{1}{2} \cdot \sum_{f=1}^{F} \sum_{f\prime=1}^{F\prime} \left(\left| TE_f - TE_{f\prime} \right| \cdot \sqrt{\left(x_{Cen_f} - x_{Cen_{f'}}\right)^2 + \left(y_{Cen_f} - y_{Cen_{f'}}\right)^2} \right) \tag{4}$$

x_{Cen_f}, y_{Cen_f}:	centroid-coordinate of facility f	TE_f:	temperature factor of facility f
$x_{Cen_{f'}}, y_{Cen_{f'}}$:	centroid-coordinate of facility f‘	$TE_{f\prime}$:	temperature factor of facility f'

Cleanliness

The functionality of the objective function is similar to the objective function of temperature. Accordingly, the cleanliness factor CE_f represents the same dependencies as the temperature factor TE_f.

$$Clean_v = \left(\frac{Cleand_v}{Cleand_{max}}\right) \cdot 100\% \tag{5}$$

$Clean_v$: objective function of the cleanliness-sensitivity in variant v
$Cleand_v$: cleanliness-sensitive distance in variant v
$Cleand_{max}$: variant v with the maximum cleanliness-sensitive distance

$$Cleand_v = \frac{1}{2} \cdot \sum_{f=1}^{F} \sum_{f'=1}^{F'} \left(\left| CE_f - CE_{f'} \right| \cdot \sqrt{\left(x_{Cen_f} - x_{Cen_{f'}}\right)^2 + \left(y_{Cen_f} - y_{Cen_{f'}}\right)^2} \right) \tag{6}$$

x_{Cen_f}, y_{Cen_f}:	centroid-coordinate of facility f	CE_f:	cleanliness factor of facility f
$x_{Cen_{f'}}, y_{Cen_{f'}}$:	centroid-coordinate of facility f'	$CE_{f\prime}$:	cleanliness factor of facility f'

3. Case Study

For the experiments, an in the literature frequently used data set was analysed [20,21,19,12]. It was first introduced by IMAM AND MIR [22]. The data set contains 20 facilities. The material flow relationships were adopted without any adjustments. For the objectives temperature and cleanliness the data set was extended by randomised attributes (see Table 1). The cell size was set to 5 metres, which corresponds to a transport route with two-way traffic and an additional pedestrian path. The shape of the factory floor area is square and has an edge length of 75 metres. Accordingly, the grid has 15x15 cells. The pick-up and drop-off points were placed in the upper left corner of the facilities. Except for the facilities that have an edge length of more than three cells. In these, pick-up and drop-off points were placed in the middle of one side. For the control parameters of the optimisation a sensitivity analysis was performed. The resulting control parameters are shown in Table 2. The two-stage optimisation required an average of about 2 hours of computing time.

Table 1: Extended facility data

facility	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
TE_f	1	-1	0	0	0	1	1	1	1	1	0	1	-1	-1	0	-1	0	0	-1	-1
CE_f	0	1	1	1	0	-1	0	0	0	0	1	-1	0	0	-1	0	-1	1	-1	-1

Table 2: Parameter setting

control parameter	first stage	second stage
TL – length of tabu list	10	7
$TS_{Iteration}$	2.000 - 10.000	100
NHS-distribution	0.8 LRS; 0.2 OAS	0.8 LRS; 0.2 OAS
$LRS_{StepSize}$	5	5
$LRS_{Iteration}$	100	100
rotation of facilities	enabled	enabled
mirroring of facilities	enabled	enabled

First, an independent optimisation was performed for each objective. Then multi-criteria optimisations were executed. If two objectives were used, the weighting factors were set to 0.5 each, i.e. equal weighting. For the optimisation of all three objectives the weighting factors were chosen as follows: material flow ω_{MF} = 0.4, temperature ω_T = 0.3, cleanliness ω_C = 0.3. The results of the experiments are shown in Table 3. Figure 3 a) shows the resulting layout regarding the mono-criteria optimisation of the material flow distance. For the mono-criteria optimisations the best result was achieved concerning the respective objective. In the multi-criteria optimisations, as previously expected, a deterioration of the results occurs, because of the contradictory objectives. When analysing the objectives, it is important to remember that the material flow distance (MF_v) must be minimised, and the temperature ($Temp_v$) and cleanliness ($Clean_v$) maximised. However, the results demonstrate that multi-criteria optimisation leads to an acceptable trade-off between different objectives. Thus, the aimed holistic approach for layout optimisation was achieved.

Table 3: Experimental results

	$MF_{v=best}$ (metres)	$Temp_{v=best}$	$Clean_{v=best}$
ω_{MF} = 1	**12,860**	5,654	5,282
ω_T = 1	32,210	**10,845**	8,222
ω_C = 1	35,170	8,975	**9,932**
ω_{MF} = 0.5, ω_T = 0.5	16,160	8,083	6,929
ω_{MF} = 0.5, ω_C = 0.5	16,200	7,595	6,760
ω_T = 0.5, ω_C = 0.5	35,950	10,753	9,451
ω_{MF} = 0.4, ω_T = 0.3, ω_C = 0.3	17,830	8,838	7,320

The results regarding the material flow distance were compared with other solution approaches [20,21,19]. The presented solution approach leads to similar good results (see Table 4). However, as the other solutions were optimised without a path structure and with pick-up and drop-off points in the centre of the facilities, the layouts had to be readjusted. A manual attempt was made to reproduce the arrangement of the facilities, considering path structure and discrete layout representation. For example, Figure 3 b) shows the adjustment concerning the result of GONÇALVES [21]. The adjustment to the path structure is expected to lead to a deterioration of the objectives. Therefore, the comparison must be examined critically.

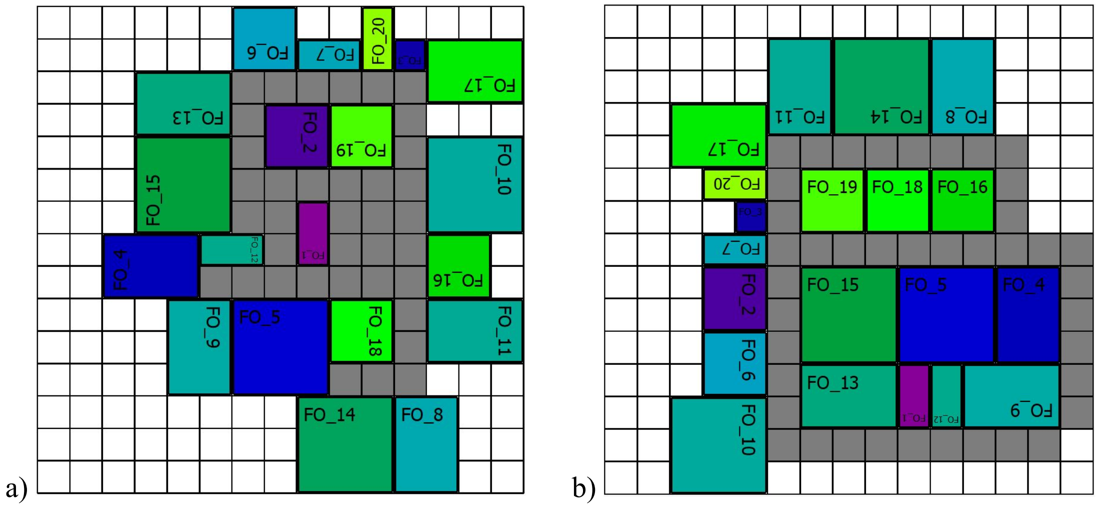


Figure 3: a) best result of MeFaP (ω_{MF} = 1), b) transmitted layout from GONÇALVES [21]

Table 4: Comparison of results with other solution approaches regarding material flow distance (MF)

$\omega_{MF} = 1$	MeFaP	[20]	[21]	[19]
$MF_{v=best}$ (metres)	12,860	21,510	16,420	19,320

4. Facility layout planning software

The previously described solution approaches were implemented with python. Additionally, a user interface was developed (Figure 4). With the resulting software, companies are able to optimise layouts independently.

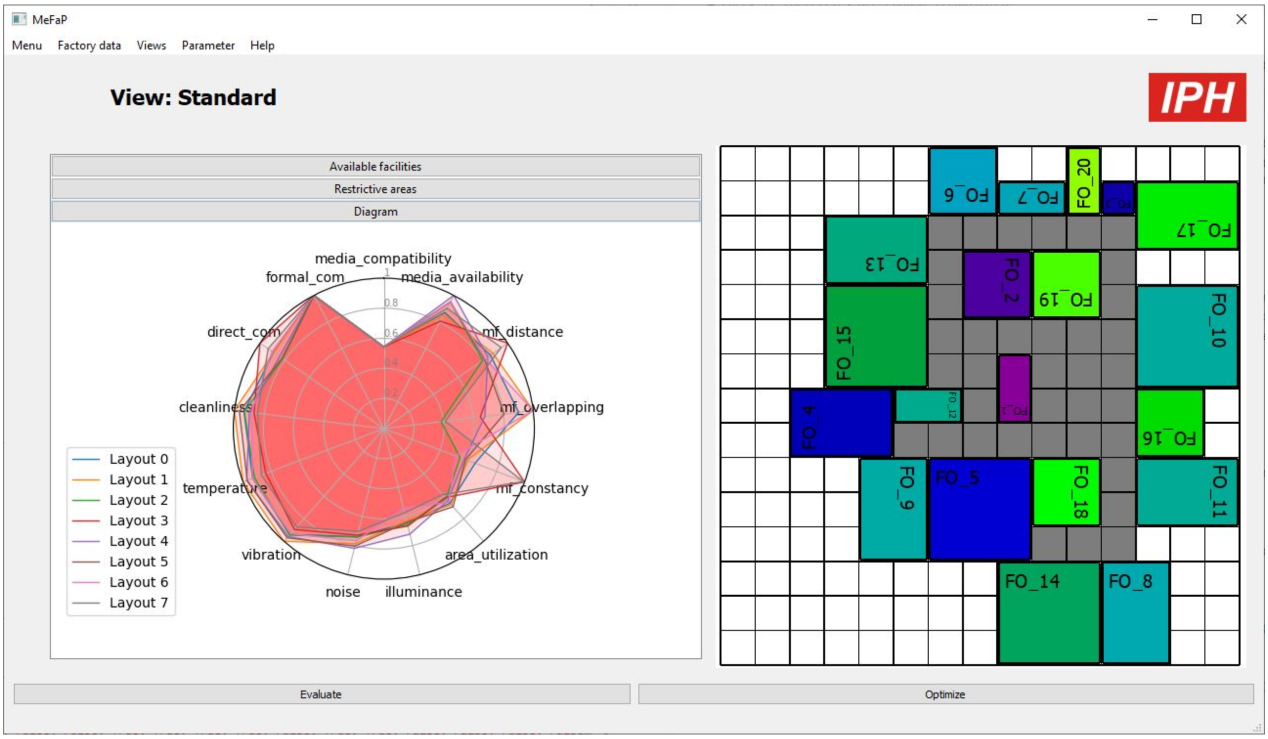


Figure 4: Example of the graphical user interface

The input data (e.g. facility data, media, material flow relationships) can be prepared in an Excel file and imported into the software. In future, it will be even possible to automatically capture facility data [23]. Factory floor space, restrictions and media availability can be planned in the software. Besides optimisation, manual planning and evaluation is also possible. In this way, existing layouts can be analysed. In addition, optimised layouts can be modified by a user and re-evaluated. The software is available as a free download (mefap.iph-hannover.de). The code is open source (gitlab.com/iph-group/lo_aif_mefap_2017/mefap).

5. Conclusion

In this paper a new solution approach and software for the facility layout problem was presented. The focus of the research project was not the development of a solution approach that is as powerful as possible, but rather to provide an easy decision support for companies. As is often the case with decision support applications for real problems, subsequent changes may be necessary, for example, to create a layout that is natural to the human eye. However, this does not mean that restrictions are not considered during the optimisation. The implemented heuristic approach is nevertheless suitable to create good layouts. The comparability with existing solution approaches could be proven. However, future improvements are possible. For example, additional solution approaches can be implemented in the software. The required computing time should be reduced by more efficient coding. The user-friendliness of the software may be improved continuously in cooperation with companies.

Acknowledgements

The IGF-Project 19666 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 Energy based on a resolution of the German Bundestag.

References

[1] Aurich, P., 2020. Schlussbericht: Entwicklung einer Methode zur quantitativen, mehrdimensionalen Fabriklayoutplanung mittels mathematischer Modellierung von fabrikplanungsrelevanten Eigenschaften: IGF-Vorhaben Nr. 19666 N. Schlussbericht, Hannover.

[2] Aurich, P., Stonis, M., Nyhuis, P., 2018. Entwicklung einer Methode zur quantitativen, mehrdimensionalen Fabriklayoutplanung. ZWF 113 (3), 117–120.

[3] Aurich, P., Stonis, M., Nyhuis, P., 2021. Software für eine automatisierte mehrdimensionale Fabriklayoutoptimierung. ZWF 116 (5), 299–302.

[4] Böning, C., 2017. Abschlussbericht: Quantitative, mehrdimensionale ad hoc Fabrikbewertung mittels mathematischer Modellierung von fabrikplanungsrelevanten Eigenschaften: IGF-Vorhaben Nr. 18111 N. Abschlussbericht, Hannover.

[5] Hosseini-Nasab, H., Fereidouni, S., Fatemi Ghomi, S.M.T., Fakhrzad, M.B., 2018. Classification of facility layout problems: A review study. The International Journal of Advanced Manufacturing Technology 94 (3), 957–977.

[6] Drira, A., Pierreval, H., Hajri-Gabouj, S., 2007. Facility layout problems: A survey. Annual Reviews in Control 31 (2), 255–267.

[7] Sharma, P., Singhal, S., 2016. A review of objectives and solution approaches for facility layout problems. International Journal of Industrial and Systems Engineering 24 (4), 469–489.

[8] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2015. Handbook Factory Planning and Design. Springer, Berlin, Heidelberg.

[9] Verein Deutscher Ingenieure, 2011. Fabrikplanung: Blatt 1: Planungsvorgehen. Beuth, Berlin. Accessed 7 May 2019, 24 pp.

[10] Liu, Q., Meller, R.D., 2007. A sequence-pair representation and MIP-model-based heuristic for the facility layout problem with rectangular departments. IIE Transactions 39 (4), 377–394.

[11] Moslemipour, G., Lee, T.S., Rilling, D., 2012. A review of intelligent approaches for designing dynamic and robust layouts in flexible manufacturing systems. The International Journal of Advanced Manufacturing Technology 60 (1-4), 11–27.

[12] Mohamadi, A., Ebrahimnejad, S., Soltani, R., Khalilzadeh, M., 2019. A New Two-Stage Approach for a Bi-Objective Facility Layout Problem Considering Input/ Output Points Under Fuzzy Environment. IEEE Access 7, 134083–134103.

[13] Bock, S., Hoberg, K., 2007. Detailed layout planning for irregularly-shaped machines with transportation path design. European Journal of Operational Research 177 (2), 693–718.

[14] Klausnitzer, A., Lasch, R., 2019. Optimal facility layout and material handling network design. Computers & Operations Research 103, 237–251.

[15] Ripon, K.S.N., Glette, K., Khan, K.N., Hovin, M., Torresen, J., 2013. Adaptive variable neighborhood search for solving multi-objective facility layout problems with unequal area facilities. Swarm and Evolutionary Computation 8, 1–12.

[16] Paes, F.G., Pessoa, A.A., Vidal, T., 2017. A hybrid genetic algorithm with decomposition phases for the Unequal Area Facility Layout Problem. European Journal of Operational Research 256 (3), 742–756.

[17] Glover, F., 1986. Future paths for integer programming and links to artificial intelligence. Computers & Operations Research 13 (5), 533–549.

[18] Hart, P., Nilsson, N., Raphael, B., 1968. A Formal Basis for the Heuristic Determination of Minimum Cost Paths. IEEE Transactions on Systems Science and Cybernetics 4 (2), 100–107.

[19] Hasda, R.K., Bhattacharjya, R.K., Bennis, F., 2017. Modified genetic algorithms for solving facility layout problems. International Journal on Interactive Design and Manufacturing 11 (3), 713–725.

[20] Derakhshan Asl, A., Wong, K.Y., Tiwari, M.K., 2016. Unequal-area stochastic facility layout problems: solutions using improved covariance matrix adaptation evolution strategy, particle swarm optimisation, and genetic algorithm. International Journal of Production Research 54 (3), 799–823.

[21] Gonçalves, J.F., Resende, M.G., 2015. A biased random-key genetic algorithm for the unequal area facility layout problem. European Journal of Operational Research 246 (1), 86–107.

[22] Imam, M.H., Mir, M., 1993. Automated layout of facilities of unequal areas. Computers & Industrial Engineering 24 (3), 355–366.

[23] Melcher, D., Küster, B., Stonis, M., Overmeyer, L., 2018. Optimierung von Fabrikplanungsprozessen durch Drohneneinsatz und automatisierte Layoutdigitalisierung.

Biography

Paul Aurich (*1991) studied industrial engineering at the Otto-von-Guericke-University Magdeburg, focused on mechanical engineering and logistics. Since November 2017 he has been working at IPH – Institut für Integrierte Produktion Hannover gGmbH as a project engineer in the field of logistics. In research and consulting projects he deals with factory planning and operations research.

Marc Speckmann (*1997) studied computer science at the Leibniz University Hannover. Since October 2019 he has been working at IPH – Institut für Integrierte Produktion Hannover gGmbH as a research assistant. In research projects he deals with factory planning and software development.

Christian Böning (*1984) has been working as project engineer in the field of logistics at IPH - Institut für Integrierte Produktion Hannover gGmbH since 2011. In research and consulting projects he deals with energy cost-oriented machine scheduling and factory planning. Since August 2013 he is head of the logistics department.

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.

2nd Conference on Production Systems and Logistics

Deep Reinforcement Learning In Production Planning And Control: A Systematic Literature Review

Marcel Panzer[1], Benedict Bender[1], Norbert Gronau[1]

[1]*Chair of Business Informatics, esp. Processes and Systems, University of Potsdam, Potsdam, Germany*

Abstract

Increasingly fast development cycles and individualized products pose major challenges for today's smart production systems in times of industry 4.0. The systems must be flexible and continuously adapt to changing conditions while still guaranteeing high throughputs and robustness against external disruptions. Deep reinforcement learning (RL) algorithms, which already reached impressive success with Google DeepMind's AlphaGo, are increasingly transferred to production systems to meet related requirements. Unlike supervised and unsupervised machine learning techniques, deep RL algorithms learn based on recently collected sensor- and process-data in direct interaction with the environment and are able to perform decisions in real-time. As such, deep RL algorithms seem promising given their potential to provide decision support in complex environments, as production systems, and simultaneously adapt to changing circumstances.

While different use-cases for deep RL emerged, a structured overview and integration of findings on their application are missing. To address this gap, this contribution provides a systematic literature review of existing deep RL applications in the field of production planning and control as well as production logistics. From a performance perspective, it became evident that deep RL can beat heuristics significantly in their overall performance and provides superior solutions to various industrial use-cases. Nevertheless, safety and reliability concerns must be overcome before the widespread use of deep RL is possible which presumes more intensive testing of deep RL in real world applications besides the already ongoing intensive simulations.

Keywords

Deep Reinforcement Learning; Machine Learning; Production Planning; Production Control; Systematic Literature Review

1. Introduction

Today's production has to cope with significantly increased complexities due to accelerating innovation cycles and increasingly individualized customer demands. Fully customized products and on-demand production impose high challenges on the associated production systems. In particular, production planning and control must be able to deal with uncertainties and constantly changing production environments [1]. Failures must be compensated quickly to enable on-time deliveries and optimize the production performance. Besides, production logistics must be able to perform the planned actions and meet the same requirements to ensure high robustness and reduce downtimes [2].

One opportunity to fulfill the demanding requirements and to keep up with product development is the application of machine learning in production systems such as (semi-)supervised, unsupervised, or reinforcement learning (RL). In contrast to (semi-)supervised and unsupervised learning, RL does not require a pre-labeled set of data and any human supervision. It is characterized in particular by its trial-and-error learning approach in direct interaction with the environment [3] and enables real-time online decision-making and an adaptive system design [4]. Especially with the success of DeepMind´s AlphaZero [5], neural network based RL received special attention which has resulted in a large number of publications in various fields and emphasized its capabilities in complex systems. However, even though [6] already emphasized the potential of general machine learning in production to improve quality and increase performances and availabilities, no focused review on research outcomes was conducted for the deployment of deep RL in production in recent years. In contrast, the fields of CPS [7] or general economics [8] among others have outlined research findings in a bundled manner and elaborated the advantages as well as major issues yet to be solved.

We intend to provide a systematic literature review of ongoing deep RL research in production planning, control, and logistics. This includes the identification of simulated and real-world implementations as well as current implementation challenges. We also want to derive possible future research directions and provide incentives to leverage the deployment of deep RL in applications that can benefit from its flexibility and adaptability. For this purpose, we intend to answer the following research questions in production planning, control and logistics:

- RQ1: What deep RL applications exist in the field of production planning, control and logistics?
- RQ2: What are existing implementation challenges of deep RL?
- RQ3: What are future research fields that need to be addressed to overcome these challenges and support implementations of deep RL in production systems?

To answer these research questions, we first give a short introduction to deep RL in Section 2, followed by the applied review methodology in Section 3. Section 4 presents the results of our review analysis (RQ1). Section 5 addresses RQ2 by outlining existing challenges and RQ3 by giving incentives for potential future research. Finally, a conclusion is given in Section 6.

2. State-of-the-art

RL is based on the agent-environment interaction loop as illustrated in Figure 1 and can be described as a sequential decision-making process. The agent performs an action and receives in turn a reward for this action and the current environmental state. With each loop and the gathered experience, the agent can adapt its behavior policy accordingly [3].

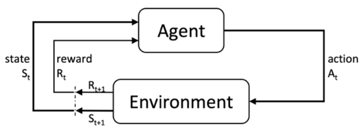


Figure 1: Agent-environment interaction loop [3]

Conventional RL methods often employ a Q-table for mapping the policy, in which recommendations for given states and the resulting actions can be retrieved. However, in high dimensional problem spaces, this leads to the curse of dimensionality and declining performances [9]. To circumvent this limitation, a neural network can be employed to map the policy. First demonstrated in 2013, such a deep RL algorithm outperformed human benchmarks in performances within the Atari environment [10]. Apart from the neural network, it is essential to distinguish between model-free and model-based algorithms. Model-based algorithms learn a general model of their environment and can make predictions about the next possible state. Model-free algorithms, on the other hand, do not learn a model of their environment, but iteratively gather experience and exploit their policy to evaluate executable actions [3]. Model-free algorithms can further be classified into value-based algorithms that require a discretization of the action space but have a better sample efficiency like the DQN, policy-based algorithms like a PPO that learn the policy directly and don´t need to evaluate actions based on Q-values like the DQN, and hybrid algorithms that try to combine both previously mentioned methods benefits [11].

A further characteristic of deep RL is its suitability for an application in decentralized and distributed multi-agent systems. Especially in the field of smart manufacturing and Industry 4.0, distributed systems can leverage the adaptability of a system and enable more robust responses against uncertainties and unforeseen events [12]. This makes RL being a promising technique to improve the performance of modern production systems.

3. Research methology

Before conducting the analysis, it is essential to establish a systematic review procedure to ensure a representative coverage of results for deep RL applications in production systems. In the further course, we follow the guidelines proposed by [13] and [14] and focus on the taxonomy as outlined in Table 1.

Table 1: Taxonomy framework

Characteristic	**Categories**			
Focus	Research outcomes	Research methods	Theories	Applications
Goal	Integration	Criticism	Central issues	
Perspective	Neutral representation	Espousal of position		
Coverage	Exhaustive	Ex. and selective	Representative	Central/pivotal
Organization	Historical	Conceptual	Methodological	
Audience	Specialized scholars	General scholars	Practitioners	General public

During the review process, we focus on research outcomes and applications of deep RL in production systems. Thereby we try to give integrative insights but also maintain a neutral position to highlight central issues that may block an implementation but also serve as research opportunities. Within our review scope, we provide a representative coverage of our chosen topic that addresses practitioners and general scholars.

For further refinement, we defined the keywords as listed in Table 2. Besides an artificial intelligence subset, a second subset describes the production domain, and a third the respective discipline.

Table 2: Defined keyword combinations

Deep RL subset			Domain subset		Discipline subset
Deep reinforcement learning OR		AND	Production OR Manufacturing OR Assembly OR Automation	AND	Planning OR Control OR Scheduling OR Dispatching OR Logistics
Reinforcement learning AND	Artificial intelligence OR Deep learning OR Machine learning				

During the review, we screened the retrieved literature from Web of Science, IEEE Xplore, and ScienceDirect (Title, abstracts and keywords, similar to [15]) according to pre-defined inclusion and exclusion criteria. We only considered English-language papers published after 2010, as deep RL and particular achievements in this field were achieved after the publication of [10] in 2013. In addition, only papers that received a peer review were included to ensure a high review quality. We included papers that focus on the impact of deep RL in production planning, control, and logistics. Purely technical papers or papers that focus on the development of algorithms were excluded. A summary of the review process is given in Figure 2. Remarkably, a large number of publications was excluded after the full-text review, since RL was considered as a machine learning technique, but did not utilize a neural network for the task completion.

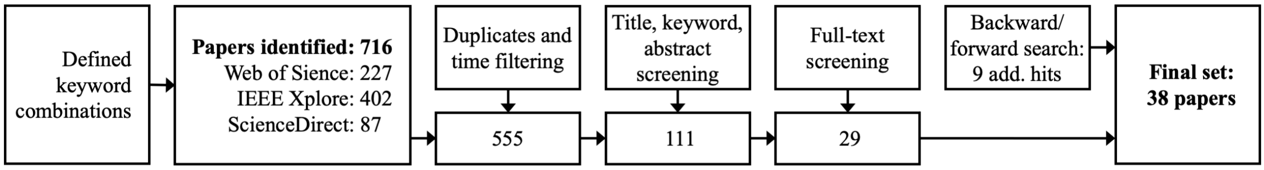
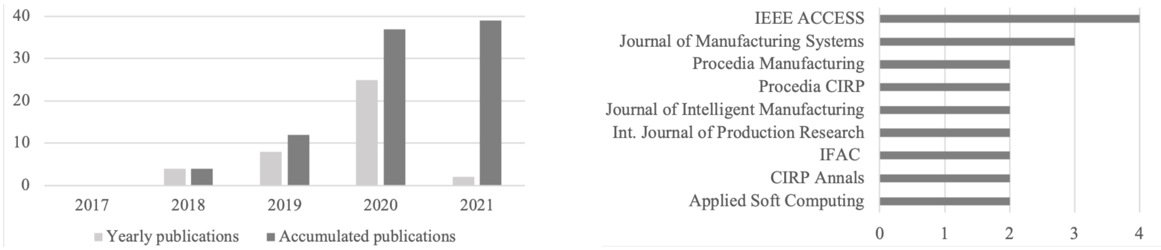


Figure 2: Review process

The distribution of publications over the years is illustrated in Figure 3. It is noticeable that no papers were published until 2017, but since 2018 there has been a significant increase, which underlines the ongoing focus on deep RL applications in current production research. On the other hand, the distribution among journals with more than one publication in Figure 4 indicates the high quality and relevance of published papers that appeared in highly recognized outlets.



Figures 3/4: Yearly publications and outlet contributions (2021 up to February)

4. Review analysis

During the in-depth analysis of the reviewed papers, production planning, as the discipline with the most publications (21), is considered first. Afterwards, due to the high overlap, production control (11) and logistics (6) are addressed in detail.

4.1 Production planning

In the field of production planning and especially scheduling, most of the reviewed papers were driven by the need to find more robust solutions that can deal with planning uncertainties and unforeseen incidents. Conventional algorithms have difficulties to cope with the dynamic production environment and often rely on human intervention or experience [16].

To deal with uncertainties, [17] increased a chemical plant's profitability by applying a policy-based algorithm for chemical production scheduling, outperforming the otherwise commonly employed MILP algorithm. Similarly, conventional algorithms in this field often have difficulties in dealing with down-times, delays, and rush orders, whereas deep RL algorithms demonstrate increased robustness. This is also evident in [18], who leveraged production scheduling to cope with highly spontaneous orders in the medical mask production during times of Covid-19. Additionally, fitted with a long-short term network, the algorithm responded more flexibly to inputs and operated faster with a reduced tardiness. Another set of papers focused on job-shop scheduling and reduced the makespan compared to FIFO, LPT/SPT, or other heuristics [19][20][16] or optimized tardiness levels, profits, and utilization rates [21]. To be more adaptive to differing problem granularities, [22] decomposed the general objective into a local and a global optimization and training problem.

In general, it is noticeable that 70% of the papers in the field of production scheduling utilized value-based algorithms that require a discretization of the action space. However, this is often feasible and can correspond to the selection of defined operations to reduce process time [23], or the selection to allocate a product to a specific machine [24]. On the other hand, a variety of inputs can be processed as demonstrated in [25] and

[26], who take Gantt diagrams for rescheduling processes as inputs. This not only led to a reduced tardiness, but also increased the flexibility of the system in handling diminishing shop-floor predictabilities.

Due to the great flexibility of the algorithm's reward function, it can be flexibly adapted and trained for other production optimization problems. To provide a brief application overview, the following Table 3 consolidates the reviewed papers within the scope of production scheduling. Even though it becomes evident that almost all of the proposed solutions outperformed conventional algorithms, they were all evaluated in simulations.

Table 3: Deep RL applications in production scheduling

Schedulings process	**Superiority to conv. methods**	**Objective**	**Source**
Chemical prod. scheduling	Superior	Maximize profits	[17]
Cloud manufacturing	Superior	Maximize utilization	[27]
Dynamic scheduling	-	Minimize completion time	[28]
	Superior	Minimize makespan	[29]
Flow shop scheduling	Superior	Reduce tardiness	[18]
Job-shop scheduling	Superior	Minimize makespan	[30]
	Superior	Minimize makespan	[19]
	-	Minimize makespan	[22]
	-	Minimize processing time	[31]
Job-shop scheduling	Superior	Minimize utilization and profits	[21]
	Superior	Minimize makespan	[20]
	Superior	Minimize makespan	[16]
Lot scheduling	Superior	Min. waiting, impr. cost-rates	[24]
Mold scheduling	Superior	Minimize processing time	[23]
Multichip production	Superior	Minimize makespan	[32]
Packaging line scheduling	Superior	Min. comp. time, energy cons.	[33]
Paint job scheduling	Superior	Minimize change-over costs	[34]
Parallel, re-entrant prod.	Comparable	High short-term return	[35]
Rescheduling	Lower tardiness	Reduce tardiness	[25]
	Lower tardiness	Reduce tardiness	[26]
Single machine scheduling	Superior	Minimize makespan, lateness	[36]

4.2 Production control

In production control, a key challenge is to compensate for sharp fluctuations in demand and breakdowns that occur at short notice to dispatch products to their respective target machines. Depending on the current state of production, orders have to be allocated to eligible machines according to their capacity, buffer levels, and further factors, while optimizing both local and global objectives [37]. To cope with the existing uncertainties, conventional methods require high computational efforts to adapt to process variations [38] or rely on single methods that do not operate optimally in each situation [39]. The static sequencing rule problem was addressed by [40] and [39], based on a situational sequencing rule selection. According to the current occupancies, machine status and others, the deep RL algorithm selected the best dispatching heuristic for the current production (such as FIFO) and significantly improved tardiness in most cases. Further approaches of adaptive job-shop scheduling were particularly investigated within the highly volatile and technically demanding wafer fabrication. By implementing deep RL driven dispatching rules, superior performances were reached compared to a variety of conventional methods, resulting in minimized time constraint violations and maintained WIP levels [41], increased machine utilization and reduced lead times [42], as well as simultaneously minimized utilization, throughput and waiting times [43]. Another approach to reduce WIP levels was proposed by [44] in production flow control. Compared to a maximum throughput method, the average WIP level could thereby be reduced by 43% with a minimal decrease in throughput (-0.2%). Further applications are listed in Table 4 and include a short-term decision-making process in mineral processing [4] which increased cumulative cash flow by 15%, a multi-agent dispatching to optimize delivery performance within the semi-conductor industry [45], and a general transfer learning supported deep RL approach in job shop processes [46]. As in scheduling, most approaches (7 out of 11) outperformed conventional algorithms, but were again solely implemented in simulations.

Table 4: Deep RL applications in production dispatching

Dispatching process	Superiority to conv. methods	Objective	Source
General job-shop disp.	Comparable	Global and local optimization	[38]
	Comparable	Minimize mean tardiness	[39]
	Superior	Minimize total tardiness	[40]
	Superior	Minimize mean lateness/tardiness	[46]
Short-term mineral flow	Superior	Optimize profits, min. target deviations	[4]
Semiconductor	Comparable	Optimize delivery performance	[45]
	Comparable	Global and local optimization	[37]
Wafer fabrication	Superior	Optimize util., TH/waiting times	[43]
	Superior	Optimize util., lead times	[42]
	Superior	Min. time constr. violations, WIP	[41]
WIP bounding	Reduces WIP	Opt. through-put and WIP trade-off	[44]

4.3 Production logistics

In [47], a real-time intralogistics solution was proposed to handle uncertainties with autonomous mobile robots (AMR). Based on the states of the individual agents, they could negotiate orders and virtually raised bids which outperformed conventional methods in terms of logistics efficiency. In a similar scenario the deep RL algorithm determined optimal target machines for the automated guided vehicles (AGV) based on job information, queue sizes, and station status which reduced lead times compared to conventional methods [48]. For the orchestration of AGVs [49] implemented a mixed rule approach. Compared to single heuristics, the makespan and delay ratio were thus reduced by approximately 10%. Besides automated vehicles, [50] proposed a deep RL algorithm for the 3-grid sorting system control to fasten up product dispatching and enable multiple sorting objectives. Other applications were a collaborative robot conveyor belt processing to fill surrounding trays [51] and a syringe filling or virtual commissioning process, among others, which outperformed human benchmarks [52]. While all compared approaches were again able to improve benchmark performances, 5 out of 6 were evaluated in simulated environments.

Table 5: Deep RL applications in production logistics

Dispatching process	Superiority to conv. methods	Objective	Source
AGV control	Superior	Min. makespan and delay ratios	[49]
	Superior	Optimize lead-times	[48]
AMR control	Superior	On-time order completion	[47]
Robot batching	-	Reach target weights and opt. filling	[51]
Syringe filling process	Above human	Min. interruptions and bad decisions	[52]
Three-grid sorting system	-	Optimize in-/outflow control	[50]

5. Implementation challenges and potential research opportunities

Despite the superior performance of deep RL, we identified the algorithm and parameter selection and optimization as well as the simulation to reality transfer as major challenges that have to be overcome to fully leverage its potential in production systems.

Beginning with the optimization, there is no guarantee for an optimal solution [17] and it is necessary to consider local as well as global optimization measures, to prevent sub-optimal solutions and decreased performances [37, 38, 45]. Besides, the choice of the algorithm, the network parameters, and further adjustment possibilities must be clarified before implementation. Regarding the algorithm selection, 22 out of 26 value-based implementations utilized a DQN, which was often inferior to enhanced versions such as a dueling or double DQN [53] and may negatively impact the performance in the particular use-case. Further challenges arise from obtaining the desired reliability and safety of the proposed solution. In production, seamless operation without incidents and maximum predictability of the system must be constantly ensured. This can only be realized by transferring the results from the simulations to reality and through subsequent intensive in-process validations. However, such a transfer to a real production system was often considered critical or required great efforts, resulting in only a few conducted real-world testings and appropriate conclusions for reality are rather hard to derive.

One way to address the above mentioned challenges in future research is, first, to test and optimize similar algorithms in parallel, which can be implemented without much efforts and contribute to a more based performance testimonial. Second, the choice of the neural network parameters can be intensively adjusted beforehand to exploit the algorithm's potential and increase its overall performance as in [22]. Further extensions such as long-short term memory or prioritized experience replay can be implemented and provide enhanced attributes for the proposed solution. Furthermore, future research should focus on an increased simulation to reality transfer. This can be accelerated significantly by multi-variable simulations which consider real-world uncertainties to minimize the existing implementation barriers. This also concerns the formulation of realistic objectives and reward functions, which do not only consider closed systems, but rather the interaction of the different actors.

To reduce general task complexities, further research can focus on a hierarchical RL frameworks, similar to the proposed rule selection framework in [39]. This circumvents the need for a single solution and distributes the varying objectives on distributed agents which are selected scenario-dependent according to pre-defined process criteria. Moreover, advanced edge functionalities can be implemented through cooperative learning and multi-agent architectures as discussed by [42] and [47]. Thus, the policy would not depend on the experience of a single agent, but would benefit from the totality of accumulated experience. The generation of a fleet intelligence would raise additional efficiencies and synergies in large and complex production systems and reduce the drawbacks of single-agent systems.

6. Conclusion

The purpose of this paper was to review existing applications of deep RL in production systems and to outline challenges and potential fields of future research. Based on a taxonomy framework, aggregated papers from three databases were narrowed to a final set of 38 papers and classified according to pre-defined criteria. It became apparent that deep RL has a broad application base in production scheduling, dispatching, and logistics, outperforming conventional algorithms in most cases and proving its ability to adapt to a wide variety of scenarios and handling production uncertainties. This not only optimized lead times, tardiness or WIP levels, but also reduced existing drawbacks of conventional methods such as high computation costs, limited adaptation capabilities, or high dependencies on human-based decisions. Nevertheless, only a few applications were assessed in reality, which makes further validation mandatory. More complex simulations that incorporate further uncertainties need to be conducted to reduce existing transfer barriers. Besides, additional consideration of optimization alternatives, such as more performant deep RL algorithms and extensions, should be considered to assess the full potential. Further research in collaborative and hierarchical multi-agent architectures and fleet intelligence approaches might also accelerate the deployment of deep RL and make it a reliable and robust optimization method for future distributed production systems.

Acknowledgements

The research presented in this paper has received funding by the Federal Ministry for Economic Affairs and Energy (BMWi), Germany (Az: 16KN086523:GeoFab – GIS2ALCM; Context-related utilization of GIS relations in life cycle management), via the VDI as project executing organization and the ZIM innovation program.

References

[1] ElMaraghy, H., AlGeddawy, T., Azab, A., and ElMaraghy, W., 2012. Change in Manufacturing – Research and Industrial Challenges, in: ElMaraghy, Enabling Manufacturing Competitiveness and Economic Sustainability, Springer Berlin Heidelberg, Berlin, Heidelberg, 2–9.

[2] Schmidtke, N., Behrendt, F., Thater, L., and Meixner, S., 2018. Technical Potentials and Challenges within Internal Logistics 4.0, in: 4th International Conference on Logistics Operations Management, IEEE, Le Havre, 1–10.

[3] Sutton, R. S., and Barto, A. G., 2017. Reinforcement Learning: An Introduction. The MIT Press, Cambridge, Massachusetts.

[4] Kumar, A., Dimitrakopoulos, R., and Maulen, M., 2020. Adaptive Self-Learning Mechanisms for Updating Short-Term Production Decisions in an Industrial Mining Complex. Journal of Intelligent Manufacturing 31 (7), 1795–1811.

[5] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., and Hassabis, D., 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. Nature 362 (6419), 1140-1144.

[6] Kang, Z., Catal, C., and Tekinerdogan, B., 2020. Machine Learning Applications in Production Lines: A Systematic Literature Review. Computers & Industrial Engineering 149, 106773.

[7] Liu, X., Xu, H., Liao, W., and Yu, W., 2019. Reinforcement Learning for Cyber-Physical Systems, in: 2019 IEEE International Conference on Industrial Internet, IEEE, Orlando, FL, USA, 318–327.

[8] Mosavi, A., Faghan, Y., Ghamisi, P., Duan, P., Ardabili, S. F., Salwana, E., and Band, S. S., 2020. Comprehensive Review of Deep Reinforcement Learning Methods and Applications in Economics. Mathematics 8 (10), 1640.

[9] Bellman, R., 1957. A Markovian Decision Process. Journal of Mathematics and Mechanics 6 (5), 679–684.

[10] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M., 2013. Playing Atari with Deep Reinforcement Learning, arXiv e-prints, arXiv:1312.5602.

[11] OpenAI, 2018. Welcome to Spinning Up in Deep RL! [Online]. Available: https://spinningup.openai.com. Accessed on: March 21 2021

[12] Rossit, D. A., Tohmé, F., and Frutos, M., 2019. Industry 4.0: Smart Scheduling. International Journal of Production Research 57 (12), 3802–3813.

[13] Tranfield, D., Denyer, D., and 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.

[14] Thomé, A. M. T., Scavarda, L. F., and Scavarda, A. J., 2016. Conducting Systematic Literature Review in Operations Management. Production Planning & Control 27 (5), 408–420.

[15] Lohmer, J., and Lasch, R., 2020. Production Planning and Scheduling in Multi-Factory Production Networks: A Systematic Literature Review, International Journal of Production Research, 1–27.

[16] Lin, C., Deng, D., Chih, Y., and Chiu, H., 2019. Smart Manufacturing Scheduling With Edge Computing Using Multiclass Deep Q Network. IEEE Transactions on Industrial Informatics 15 (7), 4276–4284.

[17] Hubbs, C. D., Li, C., Sahinidis, N. V., Grossmann, I. E., and Wassick, J. M., 2020. A Deep Reinforcement Learning Approach for Chemical Production Scheduling. Computers & Chemical Engineering 141, 106982.

[18] Wu, C.-X., Liao, M.-H., Karatas, M., Chen, S.-Y., and Zheng, Y.-J., 2020. Real-Time Neural Network Scheduling of Emergency Medical Mask Production during COVID-19. Applied Soft Computing 97, 106790.

[19] Park, J., Chun, J., Kim, S. H., Kim, Y., and Park, J., 2021. Learning to Schedule Job-Shop Problems: Representation and Policy Learning Using Graph Neural Network and Reinforcement Learning. International Journal of Production Research, 1–18.

[20] Han, B.-A., and Yang, J.-J., 2020. Research on Adaptive Job Shop Scheduling Problems Based on Dueling Double DQN. IEEE Access 8, 186474–186495.

[21] Zhou, T., Tang, D., Zhu, H., and Wang, L., 2021. Reinforcement Learning With Composite Rewards for Production Scheduling in a Smart Factory. IEEE Access 9, 752–766.

[22] Baer, S., Turner, D., Mohanty, P., Samsonov, V., Bakakeu, R., and Meisen, T., 2020. Multi Agent Deep Q-Network Approach for Online Job Shop Scheduling in Flexible Manufacturing, in: 2020 International Conference on Manufacturing System and Multiple Machines, Tokyo, Japan, 1-9.

[23] Lee, S., Cho, Y., and Lee, Y. H., 2020. Injection Mold Production Sustainable Scheduling Using Deep Reinforcement Learning. Sustainability 12 (20), 8718.

[24] Rummukainen, H., and Nurminen, J. K., 2019. Practical Reinforcement Learning - Experiences in Lot Scheduling Application. IFAC-PapersOnLine 52 (13), 1415–1420.

[25] Palombarini, J. A., and Martinez, E. C., 2018. Automatic Generation of Rescheduling Knowledge in Socio-Technical Manufacturing Systems Using Deep Reinforcement Learning, in: 2018 IEEE Biennial Congress of Argentina, IEEE, San Miguel de Tucumán, Argentina, 1–8.

[26] Palombarini, J. A., and Martínez, E. C., 2019. Closed-Loop Rescheduling Using Deep Reinforcement Learning,IFAC-PapersOnLine, 52 (1), 231–236.

[27] Zhu, H., Li, M., Tang, Y., and Sun, Y., 2020. A Deep-Reinforcement-Learning-Based Optimization Approach for Real-Time Scheduling in Cloud Manufacturing, IEEE Access, 8, 9987–9997.

[28] Zhou, L., Zhang, L., and Horn, B. K. P., 2020. Deep Reinforcement Learning-Based Dynamic Scheduling in Smart Manufacturing,Procedia CIRP, 93, 383–388.

[29] Hu, L., Liu, Z., Hu, W., Wang, Y., Tan, J., and Wu, F., 2020. Petri-Net-Based Dynamic Scheduling of Flexible Manufacturing System via Deep Reinforcement Learning with Graph Convolutional Network, Journal of Manufacturing Systems, 55, 1–14.

[30] Liu, C., Chang, C., and Tseng, C., 2020. Actor-Critic Deep Reinforcement Learning for Solving Job Shop Scheduling Problems, IEEE Access, 8, 71752–71762.

[31] Baer, S., Bakakeu, J., Meyes, R., and Meisen, T., 2019. Multi-Agent Reinforcement Learning for Job Shop Scheduling in Flexible Manufacturing Systems, in: 2019 Second International Conference on Artificial Intelligence for Industries, IEEE, Laguna Hills, CA, USA, 22–25.

[32] Park, I., Huh, J., Kim, J., and Park, J., 2020. A Reinforcement Learning Approach to Robust Scheduling of Semiconductor Manufacturing Facilities. IEEE Transactions on Automation Science and Engineering 17 (3), 1420–1431.

[33] Chen, B., Wan, J., Lan, Y., Imran, M., Li, D., and Guizani, N., 2019. Improving Cognitive Ability of Edge Intelligent IIoT through Machine Learning. IEEE Network 33 (5), 61–67.

[34] Leng, J., Jin, C., Vogl, A., and Liu, H., 2020. Deep Reinforcement Learning for a Color-Batching Resequencing Problem. Journal of Manufacturing Systems 56, 175–187.

[35] Shi, D., Fan, W., Xiao, Y., Lin, T., and Xing, C., 2020. Intelligent Scheduling of Discrete Automated Production Line via Deep Reinforcement Learning. International Journal of Production Research 58 (11), 3362–3380.

[36] Xie, S., Zhang, T., and Rose, O., 2019. Online Single Machine Scheduling Based on Simulation and Reinforcement Learning, 18. ASIM Fachtagung Simulation in Produktion und Logistik, Chemnitz, 59-68.

[37] Waschneck, B., Reichstaller, A., Belzner, L., Altenmüller, T., Bauernhansl, T., Knapp, A., and Kyek, A., 2018. Optimization of Global Production Scheduling with Deep Reinforcement Learning. Procedia CIRP 72, 1264–1269.

[38] Dittrich, M.-A., and Fohlmeister, S., 2020. Cooperative Multi-Agent System for Production Control Using Reinforcement Learning. CIRP Annals 69 (1), 389–392.

[39] Heger, J., and Voß, T., 2020. Dynamically Changing Sequencing Rules Wirth Reinforcement Learning in a Job Shop System with Stochastic Influences. Proceedings of the 2020 Winter Simulation Conference, 1608–1618.

[40] Luo, S., 2020. Dynamic Scheduling for Flexible Job Shop with New Job Insertions by Deep Reinforcement Learning. Applied Soft Computing 91, 106208.

[41] Altenmüller, T., Stüker, T., Waschneck, B., Kuhnle, A., and Lanza, G., 2020. Reinforcement Learning for an Intelligent and Autonomous Production Control of Complex Job-Shops under Time Constraints. Production Engineering 14 (3), 319–328.

[42] Stricker, N., Kuhnle, A., Sturm, R., and Friess, S., 2018. Reinforcement Learning for Adaptive Order Dispatching in the Semiconductor Industry. CIRP Annals 67 (1), 511–514.

[43] Kuhnle, A., Kaiser, J.-P., Theiß, F., Stricker, N., and Lanza, G., 2020. Designing an Adaptive Production Control System Using Reinforcement Learning. Journal of Intelligent Manufacturing 32, 855–876.

[44] Silva, T., and Azevedo, A., 2019. Production Flow Control through the Use of Reinforcement Learning. Procedia Manufacturing 38, 194–202.

[45] Waschneck, B., Reichstaller, A., Belzner, L., Altenmuller, T., Bauernhansl, T., Knapp, A., and Kyek, A., 2018. Deep Reinforcement Learning for Semiconductor Production Scheduling, in: 2018 29th Annual SEMI Advanced Semiconductor Manufacturing Conference, IEEE, Saratoga Springs, NY, USA, 301–306.

[46] Zheng, S., Gupta, C., and Serita, S., 2019. Manufacturing Dispatching Using Reinforcement and Transfer Learning, in: 2019 Joint European Conference on Machine Learning and Knowledge Discovery in Databases, 655–671.

[47] Malus, A., Kozjek, D., and Vrabič, R., 2020. Real-Time Order Dispatching for a Fleet of Autonomous Mobile Robots Using Multi-Agent Reinforcement Learning. CIRP Annals 69 (1), 397–400.

[48] Feldkamp, N., Bergmann, S., and Strassburger, S., 2020. Simulation-Based Deep Reinforcement Learning for Modular Production Systems. Proceedings of the 2020 Winter Simulation Conference, 1596–1607.

[49] Hu, H., Jia, X., He, Q., Fu, S., and Liu, K., 2020. Deep Reinforcement Learning Based AGVs Real-Time Scheduling with Mixed Rule for Flexible Shop Floor in Industry 4.0. Computers & Industrial Engineering, 149, 106749.

[50] Kim, J.-B., Choi, H.-B., Hwang, G.-Y., Kim, K., Hong, Y.-G., and Han, Y.-H., 2020. Sortation Control Using Multi-Agent Deep Reinforcement Learning in N-Grid Sortation System, Sensors, 20 (12), 3401.

[51] Hildebrand, M., Andersen, R. S., and Bøgh, S., 2020. Deep Reinforcement Learning for Robot Batching Optimization and Flow Control. Procedia Manufacturing 51, 1462–1468.

[52] Xia, K., Sacco, C., Kirkpatrick, M., Saidy, C., Nguyen, L., Kircaliali, A., and Harik, R., 2020. A Digital Twin to Train Deep Reinforcement Learning Agent for Smart Manufacturing Plants: Environment, Interfaces and Intelligence. Journal of Manufacturing Systems 58, 210–230.

[53] Wang, Z., Schaul, T., Hessel, M., Van Hasselt, H., Lanctot, M., and De Freitas, N., 2016. Dueling Network Architectures for Deep Reinforcement Learning. Proceedings of the 33rd International Conference on Machine Learning, 1995–2003.

Biography

Marcel Panzer, M.Sc. (*1994) studied mechanical engineering at the Karlsruhe Institute of Technology and has been working as a research assistant at the University of Potsdam, Chair of Business Informatics, esp. Process and Systems since 2020. His research is focused on production planning and control.

Benedict Bender, Dr. (*1989) studied business informatics at the University of Potsdam, the Humboldt University of Berlin as well as the University of St. Gallen. His research interests include Industry 4.0 and aspects of IT security and privacy. Furthermore, he deals with digital platforms and business ecosystems.

Norbert Gronau, Univ.-Prof. Dr.-Ing. (*1964) studied mechanical engineering and business administration at the Technical University of Berlin (TU). In 1994, he received his doctorate in the Department of Computer Science (TU). Up to March 2000, he was head of the teaching and research group Production-Oriented Business Information Systems at the TU Berlin. He was head of the Department of Business Information Systems at the University of Oldenburg from 2000 to 2004. Since 2004 he holds the chair of Business Informatics, Processes and Systems.

Conference On Production Systems And Logistics
CPSL 2021

2nd Conference on Production Systems and Logistics

Model-based Process Generation For Supporting Agile Shop Floor Management In SMEs

Jan Torka[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 recent years, the framework conditions on the Chinese market for machine and equipment manufacturing have changed significantly. By now, products are expected to be customized and at the same time offered at low prices with fast delivery times. Thereby, customer-specific production processes are becoming increasingly important in manufacturing. In addition, German small and medium-sized companies are confronted with a progressively turbulent business environment. For example, Chinese customers expect that orders can still be changed after they had been placed. Hence, companies must be able to identify short-term changes and individual requirements at an early stage. In order to withstand the competitive pressure, alternative courses of action must be created by generating new process variants quickly and reliably. To address this challenge, an approach for the automated generation and realization of processes based on an integrated enterprise model for supporting agile shop floor management as part of an agile process management system is being developed in the German-Chinese joint research project "*MAP*" and proposed in this paper. Starting with an introduction into the subject matter, the development gap is demonstrated and an approach as well as an early prototype is presented, followed by an outlook on further development in the research project.

Keywords

Automated process generation; Shop floor management system; Customer-specific production; Agility

1. Industrial demand

Today, many industrial companies are able to control a high number of variants and mass customization on the basis of pre-configured module structures [1]. However, customers also demand individual adaptations to standard product variants [2] [3]. Furthermore, due to an ever-increasing dynamic, product development cycles are becoming shorter. These and further factors have changed the framework in production significantly. Consequently, the complexity of production is steadily increasing [4].

Furthermore, globalization has considerably increased the importance of global trade [5]. For many decades, German companies have been active in the Chinese market. The number of German small and medium-sized companies (SMEs) in the Chinese business environment has also been growing [6]. Nevertheless, when entering a foreign market, especially the dynamic Chinese market, German companies have to adapt to the local environment and understand the local business culture [5]. That means, for example, when requesting for a quotation, Chinese customers expect suppliers to reply with an individual offer including a price and delivery time estimation in only a few days' time [7] [8]. In addition to lengthy negotiations concerning specifications, manufacturers also have to manage short-term changes to orders even though they had already been placed [9] [10] [11]. The mentioned challenges do not only apply to the Chinese market. HUANG, for example, outlines similar challenges for the U.S. market [7]. For a company to establish itself in these dynamic markets, it needs to monitor, identify and adapt to long-term trends as well as short-term changes quickly and in a targeted, well-founded manner. In this paper, following LUCZAK [12], agility is understood as the ability of a company to optimize according to different operating points in order to ensure an easy and quick change between these operating points when the scenarios occur. This ability is becoming increasingly important in process management [13] [14] [15]. To address this, an agile process management system is being developed within the framework of the German-Chinese joint research project "*Machine Learning for Agile Process Management in Machine and Equipment Manufacturing*" (MAP), which is funded by the German Federal Ministry of Education and Research (BMBF) and implemented by the Project Management Agency Karlsruhe (PTKA).

With the budatec GmbH (one of the application partners of the MAP project consortium and hereafter to be named budatec), an agile process management system is being developed for SMEs. It consists of two main interrelated components: On the one hand, a corporate environment monitoring system [16] for identifying influences at an early stage and on the other hand, an agile shop floor management system for fast and systematic reactions. In this paper, an approach and an early prototype is presented for the latter. For this purpose the application case is first described. Followed by the state of the art, where the need for research is identified. Then, in section 4, the automated generation of customer-specific production processes is shown. Based on this, the application is presented. This paper concludes with a summary and an outlook on the next steps.

2. Application case

Budatec is a leading developer and manufacturer of innovative machines and equipment for the semiconductor and solar industry. The main business areas are thermal systems and products related to electronics manufacturing. Based in Germany, the small-sized company's key strength lies in the ability to adapt its standardized products to a customer's individual requirements. Up to 80 percent of the company's products are standardized products with customer-individual adaptations and form the focus of this paper, while fully customized plants make up the remaining 20 percent. Approximately 50 percent of the application partner's machines are already destined for worldwide export. Having gained experience in Asian countries such as Japan and the Philippines, the next objective is to enter the Chinese market.

As mentioned above, due to the specifics of the Chinese market, the market entry offers both opportunities and risks [11] [17]. One of the major challenges is the time pressure in the bidding process (*short-term offer*). In order to reply to a customer request with a realistic price and delivery time estimation concerning a standard product with customized adaptation design, various factors must be included in the calculation [7] [8]. Hence, historical data of previous projects need to be collected, stored and utilized to calculate new offers. Furthermore, the company's situation in terms of, for example, liquidity needs to be monitored and simulated for the considered time period, so that decisions such as the purchase or in-house production of components can be made and the optimal offer can be designed for the customer. Another source of competitive pressure lies in the actual price level and the lead time. Therefore, orders with customer-individual adaptations need to be processed as standard orders (*processing as standard orders*), which presupposes that the time for order processing and production planning must be minimized. If the offer reaches the customer quickly enough and the rough estimations correspond to the customer's expectations, the company subsequently needs to go through several negotiation iterations (*lengthy negotiation*). Within these iterations, the specifications are being developed and the details of the offer are being negotiated. Hence, work plans, workflows and master data must be kept up to date with the negotiations. It is evident that the more transparent the production process of the specific product is, the lower are the risks of miscalculation and the better is the bargaining position of the company towards the customer. The same principle applies to changes that the customer requests after the order had already been placed (*subsequent changes*). In this case, the master data must still be adjusted shortly before or even shortly after the start of production. Classic work plan generation approaches reach their limits. Furthermore, the transparent identification and documentation of customer-requested deviations is necessary in order to enable fast and systematic processing of changes in the production process and is less prone to error than, for example, verbal assignment from the project manager to the worker on the shop floor. Finally, the commissioning of the product is often done by local partners at the customer's site (*local commissioner*), which makes the requirements for continuous documentation of the production process and a role-based view of the documentation even more significant.

In contrast to large companies, which make an effective distinction between the functions on the enterprise control level, manufacturing control level and the manufacturing level according to the automation pyramid [18], a small company as budatec needs to integrate certain functions. First and foremost, the company's key strength lies in its flexibility and customer orientation. In order to process customer-specific adaptions quickly, the order must be viewed in an integrated way. Hence, the systems for order management, customer relation management and shop floor management must be integrated as well. Added to this, the product variance of standard products is much lower compared to large companies, which lowers the need for specific information systems. Moreover, the assembly of a vacuum soldering system requires a high level of expertise and manual work, which makes many existing solutions for the automated generation of work plans for production facilities obsolete for budatec. Consequently, there is a need for an alternative shop floor management system.

The requirements derived from the challenges described above are summarized in Figure 1 and comply with the requirements for digital and intelligent shop floor management identified by RAUCH [19]. In order to react quickly and databased to short-term offers, for example, the effort for the generation of a task model for that specific order must be kept to a minimum. Additionally, in order to access existing data from previous orders and their corresponding model as a calculation basis, the maintenance effort must be kept low and the individual work steps must be documented in the first place. A further prerequisite is the access to heterogeneous data sources. The changeability of the generated model plays a major role when dealing with customer-requested changes during negotiations and occasionally also after the order had been placed.

		Challenges				
		Short-term offer	Lengthy negotiation	Processing as standard orders	Subsequent changes	Local commissioner
Requirements	Low effort for the generation of the task model for each customer individual order	x	x	x	x	x
	Low effort for model maintenance	x				
	Changeability of the generated model		x		x	
	Integration of heterogeneous data sources	x	x	x	x	x
	Integrated documentation of individual work steps	x		x	x	x

Figure 1: Requirements for the development of an agile shop floor management system, derived from the challenges of entering the Chinese market as a German SME

3. State of the art

Traditional manufacturing information technology (IT) is characterized by a hierarchical structure along the automation pyramid. This is in contrast to the service-oriented approach (SOA) which is intended to support SMEs. As mentioned in Chapter 2, it should be noted that the common components in the automation pyramid, such as the Manufacturing Execution System (MES), Enterprise Resource Planning System (ERP) and Product Lifecycle Management System (PLM), are not available in most SMEs. By outsourcing the IT infrastructure, SMEs can react more flexibly to changing market situations or customer demands [20]. In order to assist SMEs further, a concept is being developed that describes customer-specific manufacturing processes in an automated way. Considering the requirements shown in Figure 1, previous research related to automatic model generation was reviewed and compared.

All existing approaches use models, which describe the actual work process and the required resources, as a basis. These models must be able to be generated or adapted quickly and easily for new customer requests. To achieve this, the goal is to be able to compare resource capabilities automatically [21]. In order to be able to describe the capabilities in a standardized way, CUIPER uses the VDI 2680 [22] and JÄRVENPÄÄ taxonomy in comparison [21]. BACKHAUS describes a task model in his work which is used to ensure derivation and modelling of capabilities and their mapping to resource instructions and communication interfaces [23]. This approach ensures a good adaptation of the generated models. Other approaches for adjusting the model are described among others by HUCKABY and HUCKABY AND CHRISTENSEN [24] [25]. LAU goes one step further and uses an automatic comparison of capabilities to achieve a faster adaptation and generation of the models [26]. All approaches integrate a large amount of different information and documents. LAU and JÄRVENPÄÄ go the furthest in this respect, using this information for automatic capability matching [26] [21]. In order to realize a rescheduling or a documentation of the individual work steps, it must be possible to write information back into the model. LAU uses this circumstance to achieve an even utilization of machines [27].

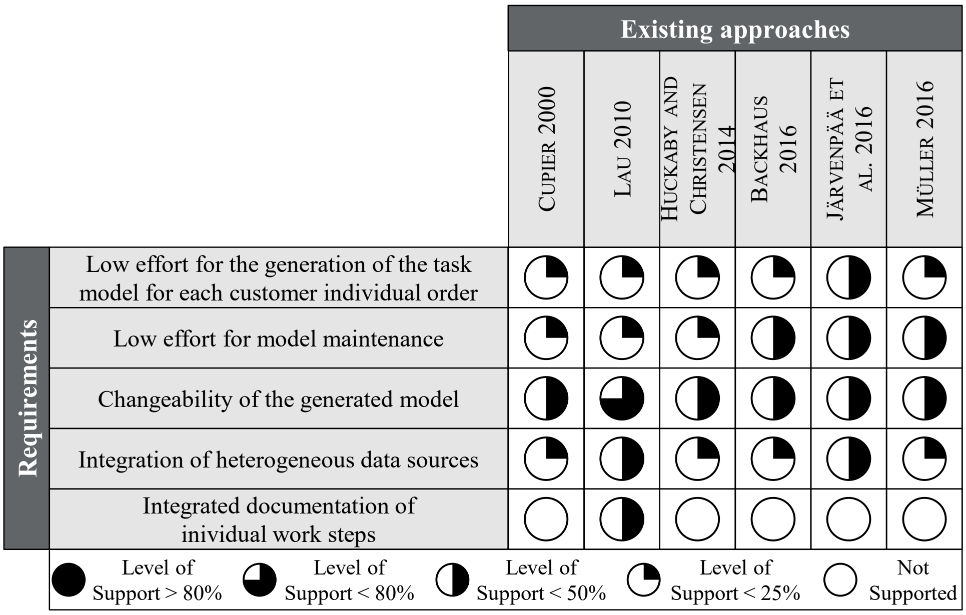


Figure 2: Comparison of existing approaches with the derived requirements

As can be seen in Figure 2, all of the approaches named above require a medium to high level of effort to create a task model. In addition, most of the steps are still manual. In CUIPER, for example, the graphically modeled description models are manually supplemented with the descriptions of the required assembly processes and the allocation of the corresponding resources [22]. Adjusting the model is possible with all of the approaches presented, but is mainly done manually as well. Changes can be implemented, however, bringing in external data sources is considered only rudimentarily by most approaches. Thus, HUCKABY ET

AL. and HUCKABY & CHRISTENSEN [24] [25] use the basic, general, atomic action considering the technical constraints and parameters. The feeding of process information back into the model is addressed by only a few approaches. The approach presented in this paper will build on the findings of the discussed scientific work. Its originality lies in the capability of taking into account order-specific product information and thereby enabling the automated generation of production processes for customer-specific adaptions to standard product variants.

4. The automated generation of customer-specific production processes

In contrast to the existing approaches for the production of product variants described above, the model-based process generation concept presented in this chapter must enable production plans to be enriched with order-specific product information – especially about the customer-individual design adaptations. In this way, order-specific production processes can be implemented. This additional information can be stored in plans and design drawings described in the following paragraphs.

The model-based approach builds on existing solutions. The Unified Service Description Language (**USDL**) [28] and the dynamic model generation, which is based on the modules of the Modular Shop floor IT [29], are worth mentioning. In the course of Industry 4.0, the linking of machines and IT as well as the amount of generated data is increasing strongly. Services are the foundation for the linking. Based on inter-connectable process modules created with the Integrated Enterprise Modelling Method (IEM) [30], a first **dynamic generated process model** was developed to describe and control a production process. This enables a flexible reconfiguration for customized products and variants as well as a faster setup of new production processes [31]. The enterprise modelling tool MO²GO [32] was used to create the described modules, which includes an information model and the actual process description [32]. The modules were developed according to the service concept.

As shown in Figure 3, the model-based approach consists of the following components, which are further explained below:

- **Business process model**, which contains the business and productions processes for the standard product variants of a company
- **Process modules**, which describe work sequences for the production of standard product variants that consist of individual work steps
- **Plans and constructing drawings**, which support the worker in production processes
- **Order management**, which contains the information of customer orders
- **Model generation system**, which generates a specific process model for an individual customer order based on the customer's requirements (standard product variant with customer-individual adaptions)
- **Specific process model for the customer order**, which contains descriptions of all work steps and the associated documents to support the worker in the production process for the individual order

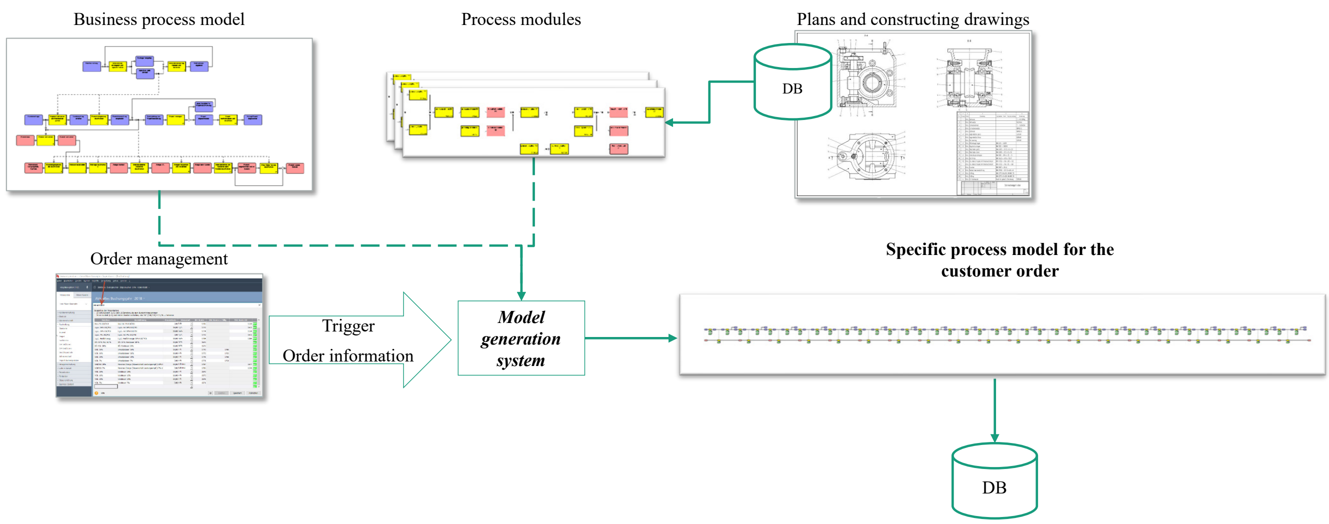


Figure 3 Concept of the model-based approach

Business process model

The business process model describes the entire value chain of the company. These workflows include purchasing, warehouse management, human resources and the actual production. It should be noted that the individual areas can differ in their level of detail. All processes that remain the same over a long time period and do not exhibit a high degree of variance can be detailed very precisely. In order to be able to follow a modular approach in automatic generation, the level of detail of the production must be selected accordingly. This means that only individual work sequences and not every single work step of the production are modelled within the generic business process model. In Figure 4, the upper halves of the process modules (illustrated as large puzzle pieces) describe the generic procedure for building a soldering oven. The detailed work sequences are represented by the small puzzle pieces in the lower half of the figure and will be referred to as "*process modules*" in the following.

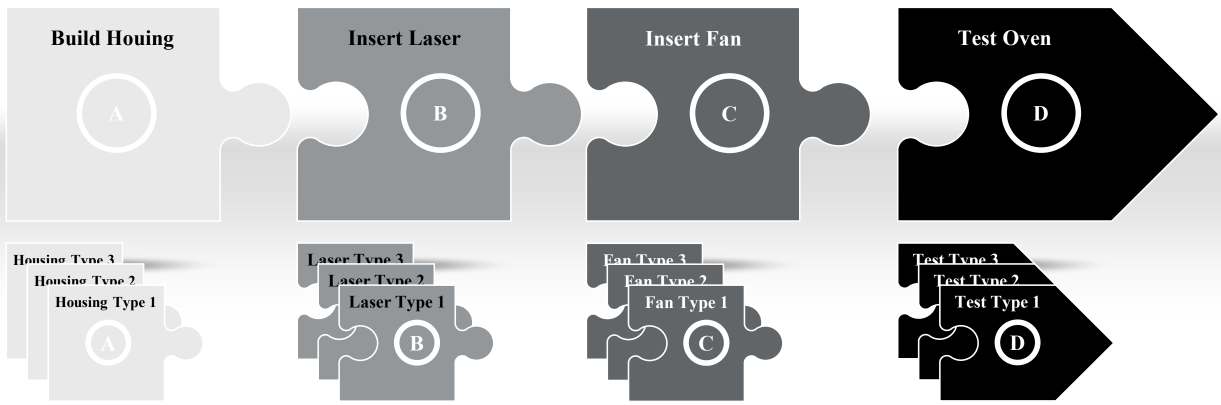


Figure 4: Concept for a generic model using the example of the production of a soldering oven

Process modules, plans and construction drawings

The exact description of the work sequences takes place in the individual process modules, illustrated as small puzzle pieces in Figure 4. Here, the necessary individual work steps are described with all aspects required for this specific task. Therefore, supporting information in the form of work instructions, plans or documentations are needed. Based on the specific customer configurations, the required documents must be adapted accordingly. If, for example, completely new production technologies are to be used for a work step, the documents must be manually checked and possibly adapted. The documents can be stored in a wide variety of locations and IT systems. Furthermore, supporting tools and necessary employees can be modelled in the modules.

Order management

The order management system handles the management of existing, past and new incoming orders. Such a system is present in most SMEs and requires a user interface. Once a new order is added to the order management system, the order information must be transferred to the downstream *model generation system*. The data to be transferred is, for example, the project number, the customer's name and the delivery date as well as the customer-individual requirements for the design adaption of the product.

Model generation system for an order-specific production model

The *model generation system* is triggered by the incoming order data and compiles the specific product model accordingly. For this purpose, the generic model from Figure 4 is taken and compared with the customer-specific product configuration and documents. Depending on the configuration, the general work sequences are now replaced with the specific work steps, and the specific production model for the customer order is generated. According to the customer order, the specific process model for the customer order can be composed of, for example, the housing type 1 with individual adjustments for increased stability (due to the increased risk of earthquakes in the customer's region), the laser type 2 and no fan. The process model created in this way is based on existing product variants but personalized by the customer-specific adjustments. The modelled individual work step descriptions contain all the necessary information to manufacture the product. In addition, the worker on the shop floor is able to save his work progress, any problems that arises or suggestions for improvement in the model for future considerations.

The requirements summarized in Figure 2 are expected to be fulfilled by the presented approach, which is currently under development. By the modular structure of the model, the effort to generate a customized model as well as the model maintenance and adaptation of existing model parts can be reduced to a minimum. This allows for quick creations of customer offers as well as for fast and coordinated reactions to short-term changes. Furthermore, new technologies can be easily mapped in the model and it is not necessary to adapt each individual model, but only the corresponding model parts. Hence, a high changeability of the model is ensured, which is especially valuable for the implementation of subsequent changes. Likewise, various data sources can be connected to support the worker with the necessary information, for example construction drawings. The model also offers the possibility to store work step related information. This makes it possible to document each individual work step. These two factors contribute to the agile implementation and tracking of short- and long-term changes to customer orders.

5. Application

The demonstrator currently consists of an early prototype in the form of a mock-up, which was initially validated with budatec. The model, whose automated creation is presented above, serves as the data basis. Figure 5a shows a general overview of all customer projects. With this, the responsible employees can quickly read the most relevant information related to an order, such as the project number, delivery date and the product to be manufactured as well as the work progress. For a better overview, completed projects are highlighted in green, whereas orange and red markings represent problematic projects.

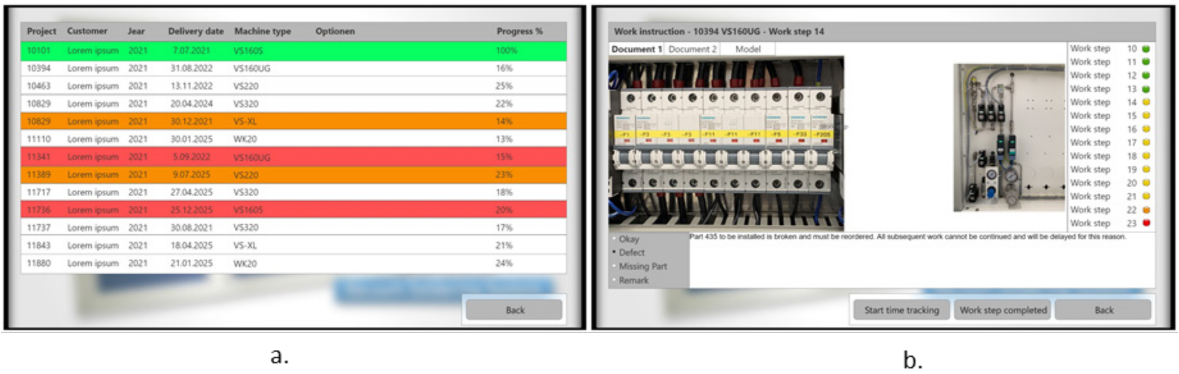


Figure 5 Mock-up of the early prototype based on a customer-specific process model

After having selected a specific project, the view shown in Figure 5b is displayed to the user. In the upper part of the screen, the order number and the product to be manufactured for this order are displayed. Detailed work instruction and design drawings can be displayed below via different tabs. On the right, the work progress and the following work steps are illustrated. In the lower area, the worker can evaluate the current work step and store additional information. In order to determine the exact duration of a work step, the worker must define the start and the end of production as well as any interruptions. The entered information is returned to the specific production model for the customer order. Afterwards, this information can be evaluated and used as a planning basis for creating new customer offers quickly and reliably in the future.

6. Summary and outlook

To establish itself in a turbulent business environment, an SME must be able to create alternative courses of actions quickly and in a target manner. However, short-term changes before and after order entry lead to significant difficulties in shop floor management. This paper derives the requirements for an agile shop floor management system from the challenges of an SME in the Chinese market and presents an approach for model-based process generation to reach these requirements. On the basis of the presented mock-up, the relevant information to be displayed on the user interface needs to be identified in the next MAP project phase as well as the corresponding data sources. Subsequently, the data needs to be imported into the process model and the user interface itself must be designed in a way that it serves as a role-based assistance system for the project manager, warehouse manager, worker and commissioner. Further considerations of the application partner include the usage of augmented reality (AR) glasses for a better and seamless display of information. Lastly, the system is planned to be validated based on a demonstrator at the application partner's site. In order to react quickly, short- and long-term changes (including for example travel restrictions, trade barriers and port shutdowns) must be identified at an early stage. For this purpose, further research will be needed in the framework of the MAP project in order to develop a methodology for classifying and evaluating influence factors as a basis for a corporate environment monitoring system.

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] Spath, D., Ganschar, O., Gerlach, S., Hämmerle, M., Krause, T., Schlund, S., 2013. Produktionsarbeit der Zukunft - Industrie 4.0. Fraunhofer-Verl., Stuttgart.

[2] McKinsey&Company, 2015. Competing for the connected customer-perspectives on the opportunities created by car connectivity and automation. https://www.mckinsey.com/~/media/mckinsey/industries/automotive%20and%20assembly/our%20insights/how%20carmakers%20can%20compete%20for%20the%20connected%20consumer/competing_for_the_connected_customer.ashx. Accessed 13 November 2018.

[3] Jovane, F., Koren, Y., Boer, C., 2003. Present and Future of Flexible Automation: Towards New Paradigms. CIRP Annals – Manufacturing Technology 1, 543–560.

[4] Lescher, M., 2008. Automatisierte Generierung von Arbeitsabläufen für den Service an Produktionssystemen. Zugl.: Aachen, Techn. Hochsch., Diss., 2008. Shaker, Aachen.

[5] Wu, J., 2008. An Analysis of Business Challenges Faced by Foreign Multinationals Operating the Chinese Market. International Journal of Business and Management 3 (12), 169–174.

[6] Bode, A., 2009. Wettbewerbsvorteile durch internationale Wertschöpfung: Eine empirische Untersuchung deutscher Unternehmen in China. Zugl.: Darmstadt, Techn. Univ., Diss., 2009, 1. Aufl. ed. Gabler, Wiesbaden.

[7] Huang, J., 2016. The challenge of multicultural management in global projects. Procedia - Social and Behavioral Sciences 226, 75–81.

[8] Huang, J., 2020. Sino-German Intercultural Management. Springer International Publishing, Cham.

[9] Faust, P., Yang, G., 2013. China Sourcing. Springer Berlin Heidelberg, Berlin, Heidelberg.

[10] ZVEI, 2014. Leitfaden-Supply-Chain-Management.

[11] Preyer, G., Krauße, R.-M., 2009. In China erfolgreich sein: Kulturunterschiede erkennen und überbrücken ; Strategien und Tipps für den Umgang mit chinesischen Geschäftspartnern, 1. Aufl. ed. Gabler, Wiesbaden.

[12] Schurig, M., 2017. Definiert: Was man unter Agilität versteht, in: Ramsauer, C., Kayser, D., Schmitz, C. (Eds.), Erfolgsfaktor Agilität. Chancen für Unternehmen in einem volatilen Marktumfeld, 1. Auflage ed. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, pp. 77–100.

[13] Fleischmann, A., Schmidt, W., Stary, C., Augl, M., 2013. Agiles Prozessmanagement mittels Subjektorientierung. HMD Praxis der Wirtschaftsinformatik 50 (2), 64–76.

[14] Schuh, G., Anderle, R., Gausemeier, J., ten Hompel, M., Wahlster, W. (Eds.), 2017. Industrie 4.0 Maturity Index: Managing the Digital Transformation of Companies (acatech STUDY). Herbert Utz Verlag, Munich.

[15] Cohn, M., Ford, D., 2003. Introducing an Agile Process to an Organization. IEEE Computer Society 36 (6), 74–78.

[16] Kohl, H., Knothe, T., Oertwig, N., Gering, P., Scholz, J.-A., 2021. Interaktives Lagebild: Ein Werkzeug für das Krisenmanagement in prozessorientierten Unternehmen. Industrie 4.0 Management 37, 37–40.

[17] Schneider, U., Fuchs, M., 2011. Internationalisierungserfolge eines KMU – Der Markteintritt der Anton Paar GmbH in China, in: Zentes, J., Swoboda, B., Morschett, D. (Eds.), Fallstudien zum internationalen management. Grundlagen - praxiserfahrungen. Gabler, pp. 381–394.

[18] VDI 5600, 2016. Manufacturing execution systems (MES).

[19] Rauch, E., Rojas, R., Dallasega, P., Matt, D.T., 2018. Smart Shopfloor Management. ZWF 113 (1-2), 17–21.

[20] Bauer, D., Stock, D., Bauernhansl, T., 2017. Movement Towards Service-orientation and App-orientation in Manufacturing IT. Procedia CIRP 62, 199–204.

[21] Järvenpää, E., Lanz, M., Tuokko, R., 2016. Application of Capability-based Adaptation Methodology for a Small-size Production System. International Journal of Manufacturing Technology and Management 30 (1/2), 67–86.

[22] Cuiper, R., 2000. Durchgängie rechnergestützte Planung und Steureung von automatisierten Montagevorgängen (143).

[23] Backhaus, J.C.S., 2016. Adaptierbares aufgabenorientiertes Programmiersystem für Montagesysteme. Utz Verlag, München.

[24] Huckaby, J., Vassos, S., Christensen, H.I. (Eds.), 2013 - 2013. Planning with a Task Modeling Framework in Manufacturing Robotics. IEEE.

[25] Huckaby, J., Vassos, S., Christensen, H.I., 2013. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems: Tokyo, Japan, November 3-7, 2013. IEEE, Piscataway, NJ.

[26] Lau, C., 2010. Methodik für eine selbstoptimierende Produktionssteuerung. Herbert Utz Verlag, München.

[27] Müller, R., Hörauf, L., Vette, M., Speicher, C., 2016. Planning and Developing Cyber-physical Assembly Systems by Connecting Virtual and Real Worlds. Procedia CIRP 52, 35–40.

[28] W3C Incubator Group, 2011. Unified Service Description Language XG Final Report. https://www.w3.org/2005/Incubator/usdl/XGR-usdl-20111027/. Accessed 23 March 2021.

[29] Knothe, T. Modulare Shopfloor IT: Komplexe IT-Architektur aus einfachen Bausteinen.

[30] Spur, G., Mertins, K., Jochem, R., 1993. Integrierte Unternehmensmodellierung, 1. Aufl. ed. Beuth, Berlin.

[31] Jaekel, F.-W., Torka, J., Eppelein, M., Schliephack, W., Knothe, T., 2018. Model Based, Modular Configuration of Cyber Physical Systems for the Information Management on Shop-Floor, in: On the move to meaningful internet systems. OTM 2017 Workshops Springer, Cham, pp. 16–25.

[32] Fraunhofer IPK, 2019. Mo²Go. www.moogo.de.

Biography

Jan Torka (*1982) is a research associate at the Fraunhofer Institute for Production Systems and Design Technology (IPK). He has previously studied Technical Computer Science at the Beuth University of Applied Sciences (M.Eng.).

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.).

2nd Conference on Production Systems and Logistics

Development Of A Systematization Of Service-Oriented Business Models For The Mechanical Engineering Industry

David Maximilian Doerr[1], Benjamin Chou[1], Thomas Bauernhansl[1]
[1]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, 70569 Stuttgart, Germany*

Abstract

Manufacturing companies, such as plant manufacturers and factory equipment suppliers, have been able to expand their traditional business model with services in recent decades, particularly around their core product. Thus, the transactional business model evolved into today's project business, where for example engineering services and after-sales services were additionally provided. Through the advancing integration of producers, customers and machines into cyber-physical systems, known as Industrie 4.0, it is now also possible to offer the core product as a service. This development is described with the help of various concepts such as service transformation, product-service system and servitization. Many of these concepts are interdependent and ambiguous. A clear scientific treatment of the concepts and especially a comparison of the concepts is therefore currently not possible.

Therefore, the purpose of this paper is to develop a systematization in the form of a morphology that describes the various concepts that constitute servitization of the core service product in the engineering industry. In doing so, the systematization is based on the four dimensions: customer, value proposition, revenue mechanics and value chain.

Keywords

Industrie 4.0, business model, mechanical engineering industry; servitization

1. Introduction

Manufacturing companies are confronted with increasingly shrinking margins and stagnating revenues. Due to globalization, this effect will continue to increase as a result of growing competition. For this reason, companies in the industrial goods sectors, such as mechanical and plant engineering, have long since refrained from competing with new rivals on price. Previous means of differentiation were a consistently high level of technological innovation, which led to ever-shorter product life cycles and a high level of investment. However, since even in the field of technological leadership, innovation cannot provide lasting protection against technologically equal competition that draws one back into price competition, companies are looking for sustainable solutions. One such solution is the development of service-oriented business models with a radical customer focus with the aim of achieving a competitive position that cannot be attacked through lasting customer loyalty [1]. For instance, understanding the customer cannot be imitated. Services have therefore been used for many years to create an increasingly individualized offer to the customer [2]. Both complementary services and the core product itself are offered as services. Companies under various names, which are partly similar but also differ in detail, provide a multitude of such offers. Research has not yet been able to provide a uniform model that systematizes these service-oriented business models. In this paper, we, therefore, want to examine the characteristics of service orientation using the example of mechanical and plant engineering in a first step. The aim of the paper is to develop a morphology of service-oriented business models. This morphology should serve to develop an understanding of the manifold characteristics of service-oriented business models.

2. Business model and service transformation in mechanical and plant engineering

Machinery and plant engineering is an important economic sector for many industrialized economies. In terms of turnover, machinery and plant engineering is the second most important industry in Germany and by far the largest industry in terms of the number of employees [3]. For many decades now, machinery and plant manufacturers have not only been pure suppliers or producers of physical products. Although products still form the core of the range of offerings, they are supplemented by services. Examples of such services are engineering, implementation and ramp-up of the machine or plant as an upstream service as well as maintenance, service and upgrades as a downstream service. This addition to the actual service increased the depth of added value and customer loyalty. This development also changed the earnings mechanics of many machine and plant manufacturers. Today, they earn less from the sale of machines, profit is rather generated over the lifetime of the machines [4]. This opens up new opportunities for machine and plant manufacturers to generate revenue beyond product sales and to further differentiate themselves. In addition, falling product margins and stagnating sales in the core business can be countered in this way.

The consistent continuation of this development with a simultaneous conceptual revolution in the sense of a business model has been finding its way into machine and plant construction for a few years now. The core product is no longer supplemented by services, but the product itself is sold as a service. In this way, the production of services or the creation of value is separated from the physical goods and only the former is sold to customers. Therefore, there is no transfer of ownership (transaction) of the machine or plant, but the service provided by the machine is paid for. An example is the company Schneider Electric, where production facilities that actually represent a high initial capital outlay are monetized with a pay-per-use model based on actual consumption. Other examples include the plant manufacturer SMS Group, which offers both entire plants and the associated equipment as an as-a-service business model. This has been accelerated by the emergence of cyber-physical systems and the networking of production facilities, known as Industrie 4.0. It is only through this digitalization and networking that automatic measurement and billing of the service provided is possible and allows these business models to emerge on a broad scale.

Machine and plant manufacturers are currently offering this new service as a complement to their classic product range. A sale of the original product as pure service is not yet taking place. The change from a producer to a service provider is also called service transformation. The service transformation is primarily about the development of new service-oriented business models that are unique and difficult to imitate and thus generate a competitive advantage [5]. This transformation process is proving increasingly difficult and time-consuming for machinery and plant manufacturers. One reason for this is that the service transformation requires high initial investments. Secondly, there is no immediate return on investment and a break-even point cannot be foreseen. Consequently, a successful service transformation is not on the horizon. The situation in which companies find themselves in this transformation process is called the service paradox.

A common cause why providers are confronted with the service paradox is that currently companies do not actively pursue systematic development of business models to achieve a self-defined goal; instead, business model development often takes place on the initiative of the customer [6]. This guarantees a corresponding demand on the market, but questions about profitability are only asked during the development of the business model or after the market launch.

The reason why there is no systematic business model development is that there is no uniform understanding of service-oriented business models. For example, the naming of the service offering varies greatly from provider to provider. Services are advertised using ambiguous and interdependent terms. On the other hand, the services offered differ when the concrete business model is analyzed in more detail. Thus, under the offered service, as described, supplementary services, as well as the core product itself, are offered as a service. These in turn differ in the details of what is offered and how a service contract is structured.

These challenges make it difficult for companies to gain an overview and thus develop an approach that ensures efficient and effective use of investments so that they can successfully go through the service transformation process.

3. Theoretical Background

The development of Industrie 4.0 as an increasingly digitalized and interconnected manufacturing industry has also increased the importance of business models. The innovation of business models describes a holistic change within the company, where various previously separate business areas work together on a business model. For example, the range of services, the way in which services are provided or the way in which payments are processed change. All these changes also require a coordinated transformation so that formerly separate business areas must network and work together. Despite the increasing importance of business models, there is no universally valid definition in the literature. Nevertheless, the common definitions all point in one direction. For example, a business model describes the way a company creates value for its customers and earns money in the process. Similar definitions can be found in Osterwalder [7], Gassmann [8] or Nagl [9].

A business model is usually described with the help of a methodology. Common methodologies are similar and differ mainly in the level of detail. For example, Osterwalder uses nine levels in the so-called Business Model Canvas to describe a business model, while Gassmann's Business Model Navigator reduces these to four elements. These elements include the value proposition (1). This describes all the benefits of a company (products and services) that are of use to customers (2) and satisfy their needs. The customer describes all customer segments to which the business model is addressed. The value chain (3) describes all processes and activities carried out by companies to realize the value proposition. This includes all resources and capabilities used along the value chain. The revenue mechanics (4) describes how the company generates revenue by fulfilling the value proposition and the costs that are incurred. The focus here is on how the company generates value.

A characteristic of digital business models that arise in the context of Industrie 4.0 is that they have a high degree of service orientation. Service transformation refers to the development of a company from a pure product manufacturer to a service provider [5]. The partial term transformation expresses that the process of change or transition is the focus of consideration [10]. Especially, in English literature, service transformation is also referred to as servitization [1], service infusion [11] or service transition [12].

The most important current research topic is the description and development of a procedure for the transformation of a manufacturing company. The number of publications on this topic has risen steadily in recent years [13].

4. Methodological approach

To obtain a comprehensive knowledge base before the actual systematization, the current state of science on business models, business model innovation and service transformation is analyzed. In addition, the status of the German mechanical and plant engineering sector will be examined in more detail. The focus is on the technologies that are being used and current challenges that must be overcome by the companies. The conceptual definition of further steps for the development of a systematization of service-oriented business models in mechanical and plant engineering is shown in Figure 1.

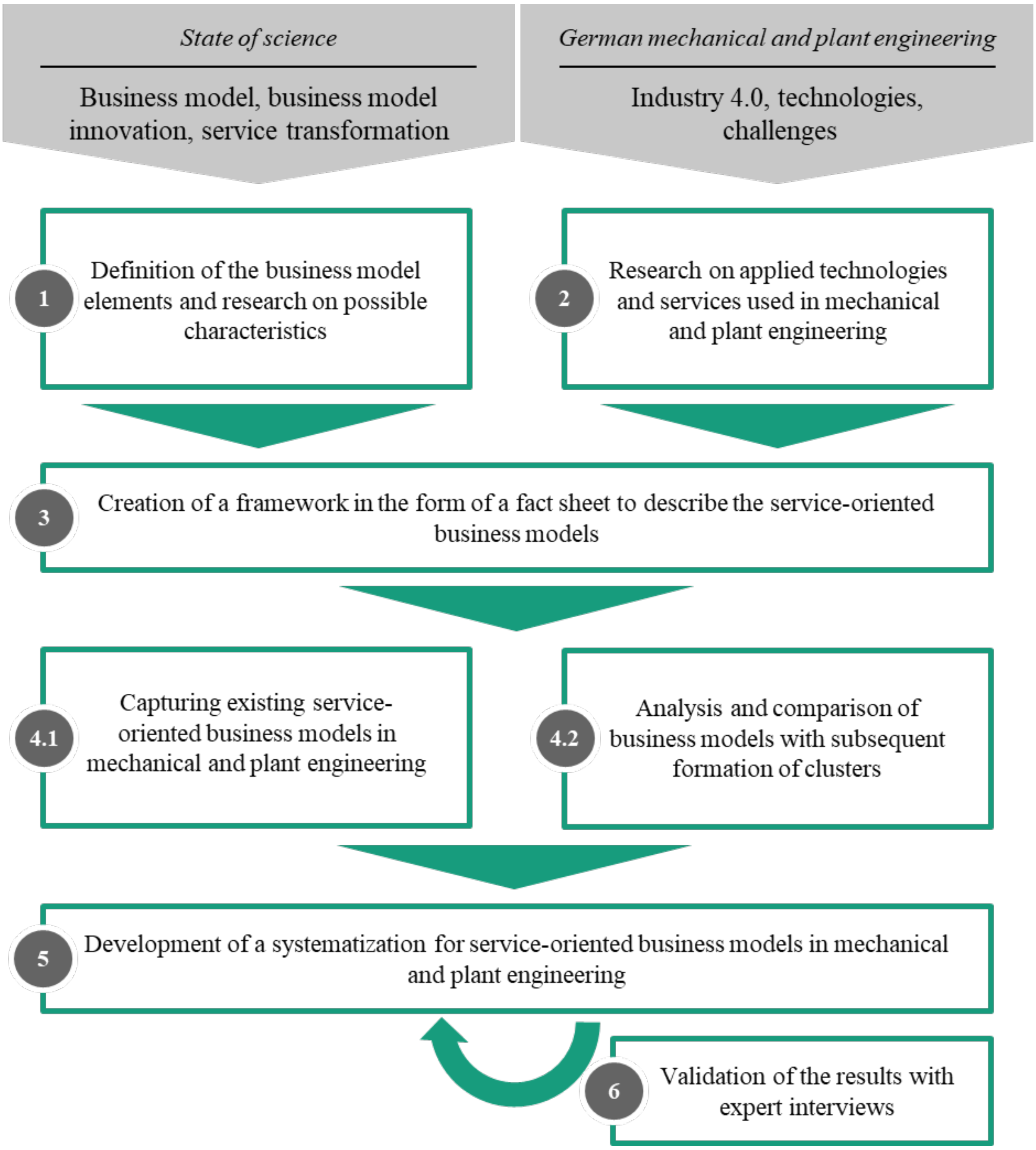


Figure 1 Conceptual approach to systematization

Research on business model elements and characteristics

In the first step, based on the findings from the literature, the business model elements are defined, which are relevant for the elaboration of this work. In addition, it is defined which conditions the business model must fulfill to be called a business model innovation. Subsequently, the literature will be researched to determine which different characteristics exist for the individual elements of the business model.

Research on technologies and services in mechanical and plant engineering

The basis for business model innovations in mechanical and plant engineering is, among other things, new technologies in connection with Industrie 4.0. In this work step, we will examine which technologies are currently offered in the mechanical and plant engineering industry and what potentials they offer for new business models. In addition, which types of services are generally offered in the mechanical and plant engineering sector, as these are core components of service-oriented business models.

Creation of a framework

With the help of the results from the first two work steps, a framework for describing service-oriented business models will be created. Descriptive characteristics that are easy to understand and measure are selected for the framework. The description characteristics are used with the determined forms of expression as a profile for the collection of practical examples. The uniform format makes it easier to evaluate and compare the recorded business models.

Capturing existing business models

Existing business models are mainly examined via an Internet search and missing information is gathered, where possible, via telephone conversations with experts. The focus of the investigation is on companies operating in the German mechanical and plant engineering market. This also includes international companies that have their headquarters outside Germany. If characteristics of the descriptive features can be found that are not yet presented in the framework, they are added.

Analysis and comparison of the business models

The recorded business models are then analyzed. An assessment is made as to whether they are business model innovations. New manifestations of the descriptive characteristics are collected and added to the previous collection. Afterward, the business models are clustered based on similarity characteristics.

Development of the systematization

The business models recorded are then analyzed. It is evaluated whether they are business model innovations. New manifestations of the descriptive characteristics are collected and added to the previous collection. Afterward, the business models are clustered based on similarity characteristics.

Validation of the results

In the final step, the systematization is validated with the help of expert interviews. During the validation, the model is presented to experts from the mechanical and plant engineering sector and the applicability and completeness of the model are checked.

5. Development of the Morphology

Through the many publications in research, the term business model has evolved over the years and many different definitions and models have emerged. Gassmann and Frankenberger's magic triangle describes the business model with four dimensions, customer value proposition, value chain and revenue mechanics. In summary, a business model is defined by who the customers are, what is sold, how it is made, and how revenue is realized [8]. The triangle shows that the cornerstones are interrelated and interdependent. According to Gassmann and Frankenberger, only when at least two of the four elements are changed, business models are innovations. For the further elaboration of the work, the magic triangle with the four elements was taken as a basis. In the next steps, the different forms of the four business model elements will be collected and compiled by researching literature and collecting practical examples. A strong focus is placed on innovative technologies that are currently being used in mechanical and plant engineering. This is because trends and technologies related to Industrie 4.0 are drivers for the development of service-oriented business model innovations. In addition, the different types of industrial services in mechanical and plant engineering are compiled and integrated into the systematization because in service-oriented business models in the manufacturing industry, classic services are usually the basis of the service offered. To capture service-oriented business models from practice and to describe their characteristics, a framework is developed in the form of a profile. The profile consists of descriptive characteristics and their forms compiled so far. The practical examples from literature and internet sources contain basic information about the business models. However, they are not comprehensive enough to serve all description features. Information about the exact revenue mechanics or profound cooperation frameworks is kept confidential by the companies and usually not available from public sources. The core of the profile consists of five selected descriptive characteristics with their respective forms of expression. The characteristics industrial services, benefit potentials, Industrie 4.0 technology, key partners and revenue sources were selected to check whether the collected expressions are comprehensive or whether further points from the practical examples need to be added. In addition, the assumption is made that the business model can be described comprehensibly and extensively with the characteristics mentioned. Based on the characteristics, the business models are later compared with each other and similarities are identified. After the creation of the profile, the next step is to collect practical examples of service-oriented business models in mechanical and plant engineering. The survey method used was against the written-postal survey, because in this classic quantitative survey the questionnaire is sent to the target group and the respondents can answer the questions in writing. With this type of data collection, one often gets a low response rate, which affects the validity of the results. Instead, practical examples were included that are described in detail in the literature, such as studies by research institutes and the website of companies from the mechanical and plant engineering sector was analyzed. Based on the available data, the fact sheet is filled in for each example. With this method, a large amount of data is obtained which can be compared through the uniform profile. Since one person collects the data, the descriptive characteristics, including forms of expression, are uniformly understood and there is little room for misunderstanding. The study focuses on machinery and plant manufacturers operating in Germany but not necessarily having their headquarters in Germany. With the assumption that companies with high sales have more financial means and resources to further develop their business models or to develop new ones, only companies with annual sales of at least 100 million euros are included in the study. In selecting the companies whose business models will be examined in more detail, one of the sources used is the European company database "Amadeus". Amadeus is a free database from Bureau van Dijk Electronic Publishing, Europe's leading provider of electronic company information. Information such as country of origin, turnover and number of employees are taken from the database if available. Otherwise, this data is taken from the company homepage. A total of 50 service-oriented business models from 45 different companies are recorded. After the service-oriented business models have been recorded using the profile, they are compared with each other based on the descriptive features and characteristics. The focus is on the identification of further characteristics of the features and commonalities of the business models. The frequency of the characteristics is counted for the practical examples, whereby multiple answers are possible. Based on the similarity characteristics, the business models are then clustered. The clusters correspond to use cases for service-oriented business models in mechanical and plant engineering., After the analysis of the recorded practical examples, the insights gained are added to the results from previous work steps. The description features used for service-oriented business models in mechanical and plant engineering with all forms of expression can be seen in full as a systematization in Figure 2.

Customer

Customer segments	Mass market	Niche market	Market segments	Diversified customer segments	Multi-sided market

Benefit

Industrial service	Assembly	Commissio-ning	Repair	Spare parts service	Maintenance	Productivity increase	
	Morder-nization	Simulation	Engineering	Software	Analyses	Procurement Service	
	Dismantling/ Disposal/ Recycling	Project-management	Planning	Consulting	Financial Service	Leasing/ Renting	
	Concessions business	General contracting	Transpor-tation services	Trainings	Teleservice	Docu-mentation	
Hybrid service bundles	Bundeling	Systems Selling	Full Service	Service Package	Product Service	Installed Base Service	Solutions
	Integrated Solutions	Eco-Efficient Producer Services	Product-Service System	Functional Sales	Functional Product	Integrated Product and Service Offering	
Degree of digitization	Product-dependent services	IT-based services	Digital services	Digitized product service systems			
Potential benefits	Increase performance	Cost reduction	Transparency of personal data	Data availability			
	Financial transparency	Increase flexibility	Risk minimization	Ressource efficiency			
Qualification of the value porposition	Function-oriented	Availability-oriented	Results-oriented				

Yield mechanics

Cost type	Fix costs	Variable costs						
Cost advantages	Economies of scale	Compound effects						
Source of income	Sales	Benefits fees	Subscrip-tion fees	Rental/ Leasing	License fees	Brokerage fees	Advertising fees	
Price mechanism	List price	Equip-ment depen-dent	Customer segment depen-dent	Quantity depen-dent	Nego-tiation	Earnings manage-ment	Real-time market	Auction

Value chain									
Key activities	Production	Problem solving	Platform/ Network						
Key resources	Financial resources	Physical resources	IT-based resources	Inventory resources	Personal resources	Structural resources	Cultural resources		
Key technology	Cyber-physical systems (CPS)	Cloud Computing	Internet of things (IoT)	Big Data	Arificial Intelligence (AI)				
	Virtual Reality/ Augmented Reality	Smart Factory	3D printing	Digital twin	Blockchain				
Key partner	Competitor	Customer	Suppliers	Banks/ Credit institions	Consulting	Research institutes	Industrial service providers		
Cooperation direction	Horizontal	Vertical	Diagonal						
Motivation for partnerships	Time advan-tages	Know-How advan-tages	Market success	System compe-tency	Cost advan-tages	Risk minimi-zation	Allocation of resources	Acquisitio nof resources	
Cooperation areas	Logis-tics	Pro-duction	Marke-ting	Sales	Service	Compa-ny infra-structure	Human resource s	R&D	Procure-ment
Cooperation intensity	Temporary	Unlimited in time							
Cooperation basic types	Joint Ventures	Strategic alliances	Company network						

Figure 2 Systematization of service-oriented business models in mechanical and plant engineering

6. Summary and Outlook

In this paper, we have developed a morphology of service-oriented business models based on previous research. To this end, we first examined the subject of study, mechanical and plant engineering, as one of the most important and strongest job-creating sectors. The service paradox was identified as the biggest problem here. Companies caught in the transformation process from a provider of physical products to a provider of services are often unable to utilize the full potential of service-oriented business models. Based on current research in the area of service-oriented business models and transformation, a current state of science and research in the area of service-oriented business models was presented. This showed that despite intensive research in the area of the transformation process, a large number of terms has been established to describe the topic. However, there was no uniform systematization and delimitation of the terms in theory. Therefore, we designed an approach in chapter 4 to develop a morphology with the help of the examination of selected practical examples. This development was then carried out in detail in chapter 5.

The benefit of this paper is, on the one hand, an initial compilation of the characteristics of service-oriented business models. In this way, companies can quickly gain an overview of the possibilities offered by service-oriented business models. For the further scientific treatment of service-oriented business models, the results of this paper offer a basis for a holistic and systematized approach to the topic.

Further research needs arise from practice. The compilation developed here is only a first step in a later process of economic assessment and selection or development of a service-oriented business model. Also, further research in similar industries like process or automotive industry makes sense, since industry specific research on business models is relatively scarce by now.

References

[1] VanDerMerwe, S., Rada, J.F., 1988. Servitization of business: Adding value by adding services. European management journal.

[2] Parida, V., Sjödin, D.R., Wincent, J., Kohtamäki, M., 2014. A Survey Study of the Transitioning towards High-value Industrial Product-services. Procedia CIRP 16, 176–180.

[3] VDMA, 2020. Maschinenbau in Zahl und Bild 2020. https://www.vdma.org/documents/105628/6872272/MbauinZuB2020.pdf/1f860d0e-9bc0-1919-639e-31da7adf53dc.

[4] Illner, B., Steynitz, G. von, 2016. Leitfaden Fit for Service. Position bestimmen, Potenziale identifizieren, Serviceerfolg steigern. https://www.vdma.org/documents/105628/878344/Fit_for_Service+Leitfaden.pdf/1d77a96a-a464-4ee4-98a1-a9d5763b565f.

[5] Bruhn, M., Hepp, M., Hadwich, K., 2015. Vom Produkthersteller zum Serviceanbieter – Vom Produkthersteller zum Serviceanbieter - Geschäftsmodelle der Servicetransformation. Marketing Review St. Gallen. http://www.bruhn-partner.com/files/2015_servicetransformation_mrsg.pdf.

[6] Emmrich, V., Bauernhansl, T., Döbele, M., 2015. Geschäftsmodell-Innovation durch Industrie 4.0. Chancen und Risiken für den Maschinen- und Anlagenbau. https://www.wieselhuber.de/migrate/attachments/Geschaeftsmodell_Industrie40-Studie_Wieselhuber.pdf.

[7] Osterwalder, A., Pigneur, Y., 2011. Business model generation: Ein Handbuch für Visionäre, Spielveränderer und Herausforderer. Campus-Verl., Frankfurt, M., New York, NY

[8] 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, XIV, 398 Seiten.

[9] Nagl, A., Bozem, K., 2018. Geschäftsmodelle 4.0: Business Model Building mit Checklisten und Fallbeispielen. Springer Gabler, Wiesbaden, XIII

[10] Ostrom, A.L., Bitner, M.J., Brown, S.W., Burkhard, K.A., Goul, M., Smith-Daniels, V., Demirkan, H., Rabinovich, E., 2010. Moving Forward and Making a Difference: Research Priorities for the Science of Service. Journal of Service Research 13 (1), 4–36.

[11] Kowalkowski, C., Persson, O., Röndell, J.G., Sörhammar, D., 2012. The co-creative practice of forming a value proposition. Journal of marketing management : MM 28 (13/14), 1553–1570.

[12] Oliva, R., Kallenberg, R., 2005. Managing the transition from products to services, in: , Operations management : a strategic approach. SAGE, London [u.a.], pp. 48–57.

[13] Baines, T., Lightfoot, H., 2013. Made to serve: How manufacturers can compete through servitization and product-service systems. Wiley, Chichester, West Sussex

Biography

David Maximilian Doerr is a research associate at the Fraunhofer Institute for Manufacturing Engineering and Automation IPA. He studied Industrial Engineering at the RWTH Aachen University. Since 2017 ,he has been coordinator for business models at Fraunhofer IPA. Since 2019, he has been working in the Corporate Strategy and Development Department. He is Deputy Manager of the Fraunhofer Project Centre for Smart Manufacturing in Shanghai.

Benjamin Chou (*1992) is an analyst at BearingPoint. He studied Industrial Engineering at the Pforzheim University of Applied Sciences (B.Eng.) and Technology Management at the University of Stuttgart (M.Sc.). During his studies, he was a student consultant at Campus Konzept Stuttgart and worked for different companies in the mechanical engineering and automotive industry. In his student research project, he was concerned with smart services in mechanical and plant engineering.

Prof. Dr.-Ing. Thomas Bauernhansl has been director of the Fraunhofer Institute for Manufacturing Engineering and Automation IPA in Stuttgart since 2011. He is also director of the Institute for Industrial Manufacturing and Management IFF at the University of Stuttgart.

Bauernhansl studied Mechanical Engineering at the RWTH Aachen University. After receiving his doctorate with distinction in 2002, he was with Freudenberg in Weinheim. His focus was on location planning, factory planning, production relocation, technology management and lean production.

The scientific focus of his research institutes is on automation, production organization, factory planning, resource efficiency in production, surface and process technology. Thomas Bauernhansl is involved in numerous advisory boards and executive committees in industry, associations, research and politics. He is the author and editor of numerous books on topics such as transformability in production, Industry 4.0 and management in production.

2nd Conference on Production Systems and Logistics

A Systematic Literature Review On The Suitability Of Cloud Migration Methods For Small And Medium-Sized Enterprises

Martin Bremer[1], Tim Walter[1], Nikita Fjodorovs[1], Katharina Schmid[1]
[1]Institute for Industrial Management, FIR at RWTH Aachen University, Campus-Boulevard 55, Aachen 52074, Germany

Abstract

The COVID-19 pandemic has shown companies that their on-premise infrastructures often reach their limits with a large number of remote accesses. The transition to cloud-based solutions could represent a more efficient alternative. However, many German companies, especially small and medium-sized enterprises (SME), are still hesitant to take this big step of transferring applications to the cloud. For this reason, this paper examines the question of whether existing migration approaches in the analysis phase fit the specific requirements of SMEs. Using a literature review methodology, we first identify and analyze determinant factors for cloud adoption in SMEs. On this basis, we analyze existing methods in the analysis phase for migrations from on-premise software to cloud solutions. We investigate whether these factors are considered in the analysis phase of the approaches and conclude their suitability for SMEs. Of the migration approaches we examined, none included all the factors we identified as relevant to SMEs. Fewer have considered all factors fully and in detail. We present the results of the literature search process in tabular form and conclude this paper with a discussion and synthesis of the literature as well as an outlook on further research fields.

Keywords

Systematic Literature Review; Cloud; Cloud-Computing, Infrastructure; Small And Medium-Sized Enterprises; SME; On-Premise; Manufacturing Companies; Business-Services

1. Introduction

Cloud computing (CC) offers great potential for small and medium-sized enterprises (SMEs), in terms of collaboration within the company and work efficiency [1], but also regarding the business domain, it can enable the realization of new products and business models [2] and unlike classical models of data processing, can meet the increasing demands on computing readiness and storage capacity [2]. It is assumed that SMEs benefit even more from cloud adoption than their larger counterparts [3], even though it must be mentioned that the decision in favor of CC requires a case-by-case consideration. The advantages for SMEs seem to contradict the fact that SMEs lack far behind larger companies in the adoption of CC [4,5].

This is a competitive disadvantage for SMEs without access to CC, which is exacerbated by the current pandemic situation. For example, insufficient bandwidth over the Internet connection for a large number of external accesses to the company's IT systems, as well as reduced data throughput rates due to VPN connections, reduced employee efficiency. Recent research has shown that adequate use of digital technologies enables better response to the COVID-19 outbreak [6,7]. Against the context of the pandemic situation, the question implied above becomes more pressing as to why the adoption of CC is not higher in SMEs. Rashmi et al. [4] mention various reasons for which SMEs actively decide against a cloud solution, such as data protection and security concerns, and internet stability. Another possible reason is that migrations to the cloud cause major challenges for SMEs, for example, they find some practices that are fitting for large companies not applicable and have difficulty hiring employees with appropriate skills [8]. For this reason, this paper evaluates existing approaches of cloud migrations in terms of their suitability for SMEs. It is intended to show whether the special determinate factors of cloud adoption in SMEs are considered in the analysis phase of cloud migration methods and to identify possible gaps. When selecting an approach for cloud migration, procedures that prove unsuitable for SMEs can be excluded accordingly. By identifying suitable migration approaches, SMEs are enabled to take advantage of the benefits of CC. These potential advantages can, for example, include a reduction in capital lockup, location-independent access, lower barriers to market entry, and access to technologies that were previously only available to large companies [2].

Existing research has already highlighted the determining factors for and impacts of the adoption of CC by SMEs. Furthermore, methods for migration were classified and compared using a characterization framework [5]. Although there have been review papers produced, we could not retrieve a secondary study that consolidates and evaluates the research on migration approaches with respect to the specific needs of SMEs in the analysis phase.

The objective of this work is to present a systematic review of literature (SLR) of existing cloud migration approaches, specifically in the analysis phase, and compare their suitability for SMEs. We focus on the business domain (e.g., 'office IT', office applications) and are therefore able to limit the variety of use-cases linked to cloud migration in general. Furthermore, we are able to take edge computing out of our scope as edge-cloud-architectures are more common in a production domain. Topics concerning 'edge-computing' will be regarded as 'on-premise' in this paper.

2. Theoretical background

2.1 Cloud computing

According to Repschläger et al. [6], CC represents 'an accumulation of services, applications, and resources that are offered to the user flexible and scalable via the internet without requiring a long-term capital commitment and IT-specific expertise'. The National Institute of Standards and Technology (NIST) further distinguishes between five characteristics (on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service) and three service models (software as a service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS)) [7]. From an organizational and corporate perspective, a distinction can be made between the deployment models 'private cloud', 'public cloud', 'hybrid cloud' and a few mixed forms [8].

2.2 Cloud migration process

A cloud migration process (CMP) is an activity or a set of activities for transferring existing enterprises' digital assets, services, IT resources or applications to a cloud. Generally, according to the three mentioned service models, there are three types of CMP – SaaS, PaaS, and IaaS migration processes. During the SaaS migration process, on-premise applications are basically replaced through a cloud-based software or service. The PaaS migration covers the process of the exchange of on-premise software operating and deployment environments to the one in a cloud. PaaS enables the creation of own applications in the cloud. The IaaS migration describes transfers of an enterprise's complete infrastructure into a cloud, including storage and network capabilities. Enterprises then use applications, servers, storage, and network resources as a service via internet connections instead of on-premise network connections [9]. IaaS migration is the most suitable opportunity for moving applications to the cloud when there is no time to adjust or redevelop the applications for a cloud. On a very general level, a cloud migration process can be divided into five phases (see Figure 1) [4].

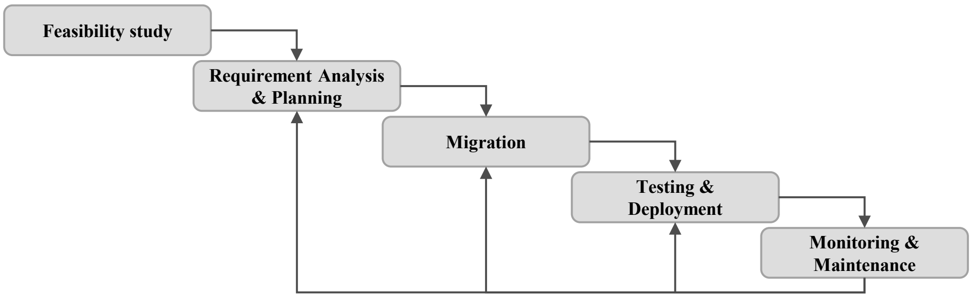


Figure 1: General five-phase model for a cloud migration process [4]

2.3 Challenges of migration to cloud-based business applications for SME

Especially small companies struggle with the selection of cloud-based solutions. Research suggests, the smaller the company, the more difficult it is for decision-makers to select the most suitable technology [10]. The risk of accidentally exposing sensitive data during the migration process is one important reason why business-critical systems are often operated on-premise [11]. SMEs also try to avoid vendor lock-in [12–19]. For this reason, it is not easy to find solutions with an appropriate balance between customization and standardization. On the other hand, cloud service providers see the necessity to provide their potential customers with a better understanding of technical and legal issues [2]. This lack of knowledge makes the assessment of service providers even more challenging from an SME's perspective. Further risks that can be identified are a high integration effort due to a lack of standards and possible misuse of data if the selection criteria are inadequate [2].

According to a survey of SMEs in Indonesia, effective cloud migration strategies and approaches can minimize the actual and perceived risks [20]. However, these may differ significantly from the actual deciding factors. To address the question of the suitability of existing migration approaches for SMEs, we examine the determining criteria for the adoption of CC in SMEs in chapter 4.2 [11]. However, these may differ significantly from the actual deciding factors. To address the question of the suitability of existing migration approaches for SMEs, we examine the determining criteria for the adoption of CC in SMEs in a later chapter.

3. Research methodology for a systematic literature review

In this paper, existing process models for cloud migrations are presented, and their suitability for small and medium-sized enterprises is investigated by applying the systematic literature review (SLR) method. An SLR is conducted to identify, evaluate and interpret available research relevant to a specific research question [21]. The selection of this method in question is based on its suitability to find gaps in existing research. We followed the procedure in [22,21], including the main phases of planning, conducting, and documenting in order to aggregate empirical evidence. The stages have been summarized and abbreviated to fit the scope of this work. Figure 2 displays our methodical approach. In the following chapters, we briefly describe our approaches during the planning, conducting, and documenting phase.

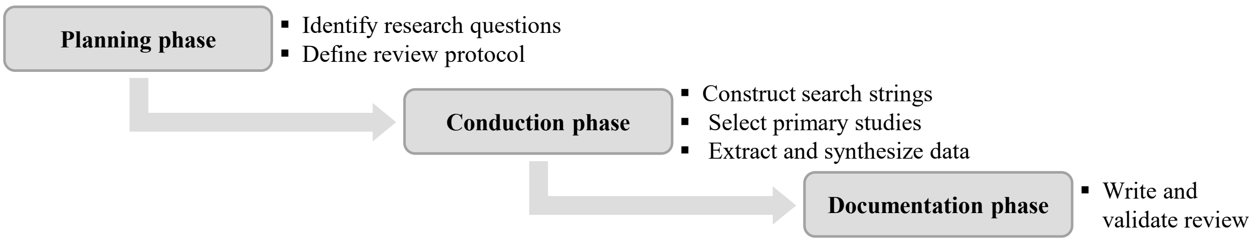


Figure 2: Our methodical approach for the SLR

3.1 Planning

In the planning phase, we identified our research questions, which are based on the objective of this work and not changed during the research process. Our identified research questions (RQ) are:

(1) *Which migration approaches do exist?*
(2) *What factors determine the decision for or against a cloud solution in SMEs?*
(3) *Are the determining factors for SMEs addressed in the migration approaches?*

Additionally, we define a review protocol in this phase. The protocol contains a precise plan for the conduction phase as well as detailed information on literature selection.

3.2 Conducting phase

The goal of the conduction phase is to find published papers (archival journals, conference proceedings, or technical reports) from the contents of predetermined electronic databases that can be used to answer the research questions. We used the following databases for our literature search: ACM Digital Library (dl.acm.org), Google scholar (scholar.google.com), IEEExplore (ieeexplore.ieee.org), Inspec (www.iee.org/Publish/INSPEC/), and ScienceDirect (www.sciencedirect.com).

We constructed strings for searching electronic databases to answer our structured research questions. Synonyms and alternative spellings were identified, and the question elements were linked with the boolean operators' OR' for synonyms and 'AND' for creating search strings. Table 1 displays the research question (RQ) and the applied search terms. For the search terms, we combined one term from each search category to form a search string. Categories were also skipped, which is represented by an epsilon (ε) within the listing of the search terms.

Table 1: Research questions and respective search terms

RQ	Research question	Search terms
1	What migration approaches do exist?	Category 1.1: method, approach, plan, technique, way, scheme, ε Category 1.2: transition, change, switch, migration Category 1.3: cloud infrastructures, cloud, on-premises to cloud
2	What factors determine the decision for or against a cloud solution in SMEs?	Category 2.1: small and medium-sized enterprises, SME, companies, Category 2.2: determinants, challenges, benefits, ε Category 2.3: transition, change, switch, migration Category 2.4: cloud, cloud infrastructures, on-premises to cloud
3	Are the determining factors for SMEs addressed in the migration approaches?	*Mainly evaluation of search results, therefore no search terms applicable*

The first selection of sources was based on a review of their title and abstract. To ensure relevance, we restricted the search to papers that were published in German and English after the year 2000. The second review of all selected studies was based on more detailed inclusion and exclusion criteria displayed in Table 2, to decide whether it can help to answer the predefined research questions.

Table 2: Inclusion and exclusion criteria for the first literature review

Inclusion criteria	Exclusion criteria
The topic of a research question is the central consideration area of the work.	The study does not cover on-premises to cloud migrations but other forms of migration such as migrations between different cloud solutions.
The work considers cloud migration in general or SaaS migrations.	The study is only published in languages other than English and German.
The study is available in full text.	Studies published before 2000.

After this step, 25 studies remained, and their citations and references were examined, as seen in Table 3.

Table 3: Results of our second literature view

Database	RQ 1 (13 studies)	RQ 2 (12 studies)
ACM Digital Library (dl.acm.org)	[23]	
Google scholar (scholar.google.com)	[9] [24] [25] [26] [27] [28]	[12] [14] [13] [29] [30] [19]
IEEExplore (ieeexplore.ieee.org)	[31] [32] [33] [34]	[35]
Springer Link (link.springer.com)	[36]	[18] [15] [17]
ScienceDirect (www.sciencedirect.com)	[40]	[37] [16]

4. Results

4.1 Migration approaches

As Zhao and Zhou [36] have already pointed out, there exist various academic research and industrial applications on legacy system migration to the cloud. Our research identified 13 relevant studies, which we further narrowed down to primary literature. Finally, eight novel frameworks, strategies, approaches, and methods for migrating to the cloud remained.

Table 4 displays existing novel migration approaches and answers our RQ1: What migration approaches do exist?

Table 4: Identified migration approaches in the primary literature

Index	Author(s)	Title	Year	Reference
1	Ahmed, Monjur; Singh, Navjot	A Framework for Strategic Cloud Migration	2019	[23]
2	Banerjee, Joydipto	Moving to the cloud: Workload migration techniques and approaches	2012	[25]
3	Khan, Nabeel; Al-Yasiri, Adil	Framework for Cloud Computing Adoption: A Roadmap for SMEs to Cloud Migration	2015	[24]
4	Nussbaumer, Nicolas; Liu, Xiaodong	Cloud Migration for SMEs in a Service Oriented Approach	2013	[27]
5	Pamami, Pooja; Jain, Aman; Sharma, Navneet	Cloud Migration Metamodel: A framework for legacy to cloud migration	2019	[33]
6	Peddigari, Bala Prasad	Unified Cloud Migration Framework — Using factory based approach	2011	[32]
7	Santikarama, Irma; Arman, Arry Akhmad	Designing enterprise architecture framework for non-cloud to cloud migration using TOGAF, CCRM, and CRMM	2016	[31]
8	Saripalli, Prasad; Pingali, Gopal	MADMAC: Multiple Attribute Decision Methodology for Adoption of Clouds	2011	[26]

4.2 Determinants of cloud computing adoption in SMEs

During the literature review, 12 papers were identified that present factors that determine the decision for or against a cloud solution in SMEs. However, factors and their influence differ from paper to paper. To identify the major influencing factors, we conducted a comparison of the named factors. We checked for each factor if it was mentioned, how detailed it was described, and what importance the authors of the paper attached to it. Importance (0 – not mentioned, 1 – low, 2 – medium, 3 – high) of respective factors was added up throughout the identified papers. As a result, we identified 15 major influencing factors. These factors do not necessarily conclude a direct advantage of CC over on-premises software. They represent what is important for SMEs when considering a migration to the cloud. The results are presented in Table 5.

Table 5: Influencing factors on the decision concerning the decision of implementing cloud solutions in SMEs (in the order of descending importance)

Factor	Literature	Score	Description
Data security concerns	[13] [14] [17] [15] [18] [35] [37] [16] [12]	21	Many SMEs are afraid of losing the data or of handing it over to a third party. However, interviewees stated that sometimes cloud services have higher data security and, in those cases, are even better than on-premise services.
Cost reduction	[13] [14] [17] [15] [18] [30] [37] [16]	21	Companies look forward to cutting their IT-related expenses through the use of cloud services because the development and maintenance of on-premise infrastructure, networks, and software may be very costly.
Reduced set-up time	[13] [17] [15] [29] [18] [30] [16]	17	Fast set-up times of services in the cloud or the opportunity to start using a service instantly is a great advantage for companies.
Dependency on provider	[13] [14] [17] [15] [18] [19] [16] [12]	16	High dependency on the provider of cloud services discourages companies from using the cloud services.
Scalability	[13] [14] [17] [29] [18] [30] [12]	16	The possibility to increase cloud resources, storage, or functions is seen positively by enterprises.
Flexibility in regard to location and devices	[13] [14] [17] [15] [18] [16] [12]	15	Cloud services allow remote collaboration as well as the use of different devices and operating systems and, in such a manner, encourage companies to use cloud services.
Flexibility in regard to used services/apps	[13] [14] [17] [15] [29] [18] [16]	14	The majority of papers state that the possibility to add or remove functions depending on the enterprise's needs persuades SMEs to choose cloud services.
Focus on the main business	[13] [14] [17] [15] [29] [30]	13	Through outsourcing of IT operations, enterprises can focus on their main business and activities.
Easiness of use	[13] [14] [17] [29] [18] [30] [12]	13	Better user experience and user interface are reasons for enterprises to choose cloud services.
Dependency on the internet and its quality	[13] [17] [15] [18] [16] [12]	13	Enterprises hesitate to migrate fully to the cloud because of possible problems with access to the data or inability to work in case of a poor internet connection.
Transparency in regard to service	[13] [14] [17] [15] [19]	10	If it is not clear how a cloud service works, what are its privacy policies, who has access to the data, and where the servers are located, enterprises will hesitate to choose the service.
Organizational size and structure	[13] [14] [17] [37]	10	The smaller the company and the flatter its hierarchy structure, the easier and more likely it is to migrate to cloud services.
Technical knowledge / understanding of CC	[13] [14] [35] [37]	9	Deeper technical knowledge of CC has a positive impact on the decision to use cloud services.
Competitive pressure	[13] [14] [17] [35] [37]	9	Competing organizations adopting cloud and increasing adoption trends across industry sectors influence the decision to use cloud services.
Easy possibility to test services/apps before purchase	[13] [17] [29]	8	Cloud services allow to try new services before their actual purchase and thus encourage enterprises to use cloud services.

4.3 Synthesis of the literature of existing methods and approaches for migration

Although the literature examined explicitly addresses SMEs, in part because of the search terms chosen, none considers all of the factors identified. Depending on the objective of the literature analyzed, only very few of the factors were even considered in the critical analysis phase. Moreover, very few authors made a distinction with regard to IaaS, PaaS, and SaaS. Table 6 summarizes the results regarding research question 3: 'Are the determining factors for SMEs addressed in the migration approaches?'. In order to evaluate to what extend do the identified address the influencing factors, the following evaluation logic was used: 4 – if an approach provides a detailed guide on how to address a factor; 3 – if an approach proposes a concept to address a factor; 2 – if an approach describes general opportunities to address a factor; 1 – if a factor was mentioned in the approach.

Table 6: Analysis of the identified migration approaches regarding the identified influencing factors
(0: not at all, 1: partly, 2: mainly, 3: mostly, 4: fully)

Factor	[23]	[25]	[24]	[27]	[33]	[32]	[31]	[26]
Data security concerns	2	2	1	4	0	1	2	4
Cost reduction	1	1	4	4	2	2	2	4
Reduced set-up time	0	0	1	3	0	1	0	3
Dependency on provider	3	2	2	0	1	0	2	1
Scalability	1	1	1	4	0	2	0	2
Flexibility regarding location and devices	1	0	2	2	1	2	3	2
Flexibility regarding used services/apps	0	0	3	4	1	2	2	3
Focus on the main business	0	2	1	0	0	0	0	1
Easiness of use	4	0	1	4	0	0	0	4
Dependency on internet and its quality	1	0	1	4	1	0	2	4
Transparency in regard to service	0	0	1	0	0	0	0	0
Organizational size and structure	4	2	0	1	4	0	4	2
Technical knowledge/understanding of CC	3	2	1	2	2	0	2	0
Competitive pressure	0	0	0	0	0	0	1	0
Easy possibility to test services/apps before purchase	0	3	3	0	0	0	0	0

The table shows that no migration approach considers all factors. There are at least two factors not considered (0) in all approaches. The migration approach with the highest coverage, considers 6 out of 15 factors completely and in detail (4). Most of the analyzed literature did not consider a fair amount of the identified factors and were either too strategic or only addressing specific factors. In addition, the analysis phase can identify potential that could be realized through a change in processes or service utilization. The factors targeting this are insufficiently considered in the literature analyzed. In conclusion, the identified methods and approaches did not meet and fit the requirements of SMEs.

5. Discussion, conclusion, and outlook

This paper presented the results of SLR of research on CC and SMEs from 2000 to 2021. We reviewed a total of 25 scientific studies to identify cloud migration approaches in scientific papers, analyze what factors determine the decision for or against a cloud solution in SMEs, and subsequently investigate if the factors are addressed in the migration approaches. As presented in chapter 4.3, the identified migration approaches do not consider important factors and may therefore not be applied by SMEs.

In conclusion, this systematic literature review underlines the importance of continued research in this area to enable SMEs to utilize the potential of cloud-based services fully. In order to move away from existing on-premise-based infrastructures, this paper shows that a variety of factors must be taken into consideration. A suitable and holistic approach that particularly considers the possibilities of SaaS for SMEs could not be identified. From the authors' point of view, there is a need to give greater consideration to the resulting advantages and develop suitable process models that take into account all the relevant influencing factors.

Acknowledgements

The research in this paper is part of the project LegitimiseIT, funded by the German Federal Ministry for Economic Affairs and Energy (project number 05339/19 N).

References

[1] Leeser, D.C., 2020. Cloud-Computing als Herausforderung für KMU, in: Leeser, D.C. (Ed.), Digitalisierung in KMU kompakt. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 113–134.

[2] Hentschel, R., Leyh, C., 2018. Cloud Computing: Status quo, aktuelle Entwicklungen und Herausforderungen, in: Reinheimer, S. (Ed.), Cloud Computing. Springer Fachmedien Wiesbaden, Wiesbaden, pp. 3–20.

[3] Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., Ghalsasi, A., 2011. Cloud computing — The business perspective. Decision Support Systems 51 (1), 176–189.

[4] Rashmi, R., Mehfuz, S., Sahoo, G., 2012. A five-phased approach for the cloud migration. International Journal of Emerging Technology and Advanced Engineering 2 (4), 286–291.

[5] Jamshidi, P., Ahmad, A., Pahl, C., 2013. Cloud Migration Research: A Systematic Review. IEEE Trans. Cloud Comput. 1 (2), 142–157.

[6] Repschläger, J., Pannicke, D., Zarnekow, R., 2010. Cloud Computing: Definitionen, Geschäftsmodelle und Entwicklungspotenziale. HMD 47 (5), 6–15.

[7] Mell, P.M., Grance, T., 2011. The NIST definition of cloud computing, Gaithersburg, MD.

[8] BITKOM, 2010. Cloud Computing – Was Entscheider wissen müssen. Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V., Berlin.

[9] Pahl, C., Xiong, H., 2013 - 2013. Migration to PaaS clouds - Migration process and architectural concerns, in: 2013 IEEE 7th International Symposium on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems. 2013 IEEE 7th International Symposium on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems (MESOCA), Eindhoven, Netherlands. 23.09.2013 - 23.09.2013. IEEE, pp. 86–91.

[10] Schröder, C., 2016. Herausforderungen von Industrie 4.0 für den Mittelstand, Bonn, 28 pp.

[11] Pahl, C., Xiong, H., Walshe, R., 2013. A Comparison of On-Premise to Cloud Migration Approaches, in: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M.Y., Weikum, G., Lau, K.-K., Lamersdorf, W., Pimentel, E. (Eds.), Service-Oriented and Cloud Computing, vol. 8135. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 212–226.

[12] Abdollahzadegan, A., Hussin, Ab Razak Che, Gohary, M.M., Amini, M., 2013. The Organizational Critical Success Factors for Adopting Cloud Computing in SMEs. Journal of Information Systems Research and Innovation (JISRI) (Volume 4, Issue 1), pp. 67-74.

[13] Deil, R., Brune, P. (Eds.), 2017. Cloudy with a Chance of Usage? – Towards a Model of Cloud Computing Adoption in German SME.

[14] Friedrich-Baasner, G., Fischer, M., Winkelmann, A., 2018. Cloud Computing in SMEs: A Qualitative Approach to Identify and Evaluate Influential Factors, in: Proceedings of the 51st Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences. Hawaii International Conference on System Sciences.

[15] Hentschel, R., Leyh, C., Egner, M., 2020. Motivationsfaktoren für oder gegen einen Einsatz von Cloud-Lösungen in Kleinstunternehmen. HMD 57 (5), 961–975.

[16] Hsu, P.-F., Ray, S., Li-Hsieh, Y.-Y., 2014. Examining cloud computing adoption intention, pricing mechanism, and deployment model. International Journal of Information Management 34 (4), 474–488.

[17] Karunagaran, S., Mathew, S.K., Lehner, F., 2019. Differential cloud adoption: A comparative case study of large enterprises and SMEs in Germany. Inf Syst Front 21 (4), 861–875.

[18] Leeser, D.C., 2020. Lokale Herausforderungen durch Digitalisierung, in: Leeser, D.C. (Ed.), Digitalisierung in KMU kompakt. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 63–111.

[19] Wu, Y., Cegielski, C.G., Hazen, B.T., Hall, D.J., 2013. Cloud Computing in Support of Supply Chain Information System Infrastructure: Understanding When to go to the Cloud. J Supply Chain Manag 49 (3), 25–41.

[20] Widyastuti, D., Irwansyah, M. Benefits And Challenges Of Cloud Computing Technology Adoption In Small And Medium Enterprises (SMEs), in: , Bandung Creative Movement (BCM) Journal, Vol 4, No 1.

[21] Kitchenham, B., Charters, S., 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering. Keele University and Durham University Joint Report, Durham.

[22] Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M., Khalil, M., 2007. Lessons from applying the systematic literature review process within the software engineering domain. Journal of Systems and Software 80 (4), 571–583.

[23] Ahmed, M., Singh, N., 2019. A Framework for Strategic Cloud Migration, in: Proceedings of the 2019 5th International Conference on Computing and Artificial Intelligence - ICCAI '19. the 2019 5th International Conference, Bali, Indonesia. 19.04.2019 - 22.04.2019. ACM Press, New York, New York, USA, pp. 160–163.

[24] Khan, N., Al-Yasiri, A., 2015. Framework for Cloud Computing Adoption: A Roadmap for Smes to Cloud Migration. IJCCSA 5 (5/6), 1–15.

[25] Banerjee, J., 2012 - 2012. Moving to the cloud: Workload migration techniques and approaches, in: 2012 19th International Conference on High Performance Computing. 2012 19th International Conference on High Performance Computing (HiPC), Pune, India. 18.12.2012 - 22.12.2012. IEEE, pp. 1–6.

[26] Saripalli, P., Pingali, G., 2011 - 2011. MADMAC: Multiple Attribute Decision Methodology for Adoption of Clouds, in: 2011 IEEE 4th International Conference on Cloud Computing. 2011 IEEE 4th International Conference on Cloud Computing (CLOUD), Washington, DC, USA. 04.07.2011 - 09.07.2011. IEEE, pp. 316–323.

[27] Nussbaumer, N., Liu, X., 2013 - 2013. Cloud Migration for SMEs in a Service Oriented Approach, in: 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops. 2013 IEEE 37th International Computer Software and Applications Conference Workshops (COMPSACW), Japan. 22.07.2013 - 26.07.2013. IEEE, pp. 457–462.

[28] Balobaid, A., Debnath, D., 2020 - 2020. An Effective Approach to Cloud Migration for Small and Medium Enterprises (SMEs), in: 2020 IEEE International Conference on Smart Cloud (SmartCloud). 2020 IEEE International Conference on Smart Cloud (SmartCloud), Washington DC, WA, USA. 06.11.2020 - 08.11.2020. IEEE, pp. 7–12.

[29] Haselmann, T., Röpke, C., Vossen, G., 2011. Empirische Bestandsaufnahme des Software-as- a-Service-Einsatzes in kleinen und mittleren Unternehmen. Arbeitsberichte des Instituts für Wirtschaftsinformatik (131).

[30] Raut, R.D., Gardas, B.B., Narkhede, B.E., Narwane, V.S., 2019. To investigate the determinants of cloud computing adoption in the manufacturing micro, small and medium enterprises. BIJ 26 (3), 990–1019.

[31] Santikarama, I., Arman, A.A., 2016 - 2016. Designing enterprise architecture framework for non-cloud to cloud migration using TOGAF, CCRM, and CRMM, in: 2016 International Conference on ICT For Smart Society (ICISS). 2016 International Conference on ICT For Smart Society (ICISS), Surabaya, Indonesia. 20.07.2016 - 21.07.2016. IEEE, pp. 32–37.

[32] Peddigari, B.P., 2011 - 2011. Unified Cloud Migration Framework — Using factory based approach, in: 2011 Annual IEEE India Conference. 2011 Annual IEEE India Conference (INDICON), Hyderabad, India. 16.12.2011 - 18.12.2011. IEEE, pp. 1–5.

[33] Pamami, P., Jain, A., Sharma, N., 2019 - 2019. Cloud Migration Metamodel : A framework for legacy to cloud migration, in: 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India. 10.01.2019 - 11.01.2019. IEEE, pp. 43–50.

[34] Ahmad, N., Naveed, Q.N., Hoda, N., 2018 - 2018. Strategy and procedures for Migration to the Cloud Computing, in: 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS). 2018 IEEE 5th International Conference on Engineering Technologies and Applied Sciences (ICETAS), Bangkok, Thailand. 22.11.2018 - 23.11.2018. IEEE, pp. 1–5.

[35] Borgman, H.P., Bahli, B., Heier, H., Schewski, F., 2013 - 2013. Cloudrise: Exploring Cloud Computing Adoption and Governance with the TOE Framework, in: 2013 46th Hawaii International Conference on System Sciences. 2013 46th Hawaii International Conference on System Sciences (HICSS), Wailea, HI, USA. 07.01.2013 - 10.01.2013. IEEE, pp. 4425–4435.

[36] Zhao, J.-F., Zhou, J.-T., 2014. Strategies and Methods for Cloud Migration. Int. J. Autom. Comput. 11 (2), 143–152.

[37] Oliveira, T., Thomas, M., Espadanal, M., 2014. Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. Information & Management 51 (5), 497–510.

Biography

Martin Bremer (*1989) started working as a project manager at FIR at RWTH Aachen University in 2017. In his current position within the technical group IT complexity management, he specialized within IT architecture management, the development of transformation strategies, enterprise cloud transformations and Business-IT-Alignment.

Tim B. Walter (*1993) has been a scientific researcher and project manager of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2020. In his current position as part of the Information Management Division, he supports companies in various industries in the design and implementation of the information logistics of business processes. He also participates in different research projects.

Nikita Fjodorovs (*1994) studied Business Administration and Engineering specializing in Industrial Engineering and Operations Management at RWTH Aachen University and Automotive Engineering at Tsinghua University as a part of a double degree program. He was also a graduate student research assistant at FIR at RWTH Aachen University. After his graduation in 2021 he started working as a project manager at FIR at RWTH Aachen in the division Production Management.

Katharina Schmid (*1992) studied Business Administration at the DHBW Ravensburg specializing on logistics and controlling. After graduation, she worked as a project manager for an international automotive supplier in supply chain projects. She deepened her knowledge in electronic processes with a study of computer science at the RWTH Aachen. She joined FIR at RWTH in 2019 in the Information Management Division.

2nd Conference on Production Systems and Logistics

Analysis of the Current Situation on Automation and Digitalization in Moroccan Industry

Johannes be Isa[1], Lukas Merkel[1], Nicolas Brixle[1], John Ulrich Fimpel[2], Ardin Djalali[3]

[1]*apt advanced production training GmbH, Augsburg, Germany*
[2]*Zentralverband Elektrontechnik- und Elektronikindustrie e.V., Frankfurt, Germany*
[3]*Steinbeis School of International Business and Entrepreneurship (SIBE) GmbH, Herrenberg, Germany*

Abstract

Morocco is establishing itself as a key industrial location and logistical hub for the African continent. The Competence Center on Automation (CCoA), funded by the Federal Ministry for Economic Cooperation and Development, BMZ has the vision to increase economic growth and to create new jobs through knowledge transfer on automation and digitalization technology in key industrial sectors. This paper shows the approach and the results of an analysis based on quantitative and qualitative interviews. The goal of this research is to analyze the degree on automation and digitalization in the selected Moroccan industries. It is based on quantitative and qualitative interviews. The target group of the pilot phase includes automotive and food industry, universities, vocational schools and political stakeholders. The results show several differences regarding technology use and training between the automotive and food industry, as well as between local and international companies in Morocco. These results will be used to develop a digitalization roadmap, tailor made for the Moroccan industry, to make the country even more competitive and to establish international joint ventures on the long term.

Keywords

Automation, Digitalization; Global Manufacturing, International Business, Knowledge Transfer

1. Motivation

The successful transformation from a developing country to an industrialized nation requires international business activities [1]. To remain an industrialized nation, it is also necessary to work on relationships at the global level. A good way to put relationship work into practice is joint success through knowledge transfer. Another long-term step to ensure common success are industrial clusters, joint cooperations and legally joint ventures. Building long-term, sustainable relationships and creating high-quality jobs in Morocco is the goal of the Competence Center on Automation (CCoA). Automation and digitalization will play a major role in the global manufacturing industry in the future [2]. Furthermore, the strategic orientation of international production networks is fundamental, for which knowledge of the interdependencies between network and factory level is essential [3]. Dynamic production networks can be a solution for managing current and future challenges such as high volatility in demand and the trend towards customized products [4]. Facing these challenges requires a transformation with the intelligent use of automation and digitalization solutions to so called cyber-physical systems which represent the core of smart factories [5]. This, in turn, demands high-quality training and further education for specialists and managers.

2. Related work

The challenges of implementing latest automation and digitalization applications in Tunisian small and medium-sized enterprises (SMEs) were analyzed in the work of Ben Hadj Hassine 2021 [6]. Therefore, an exploratory qualitative approach using a questionnaires and individual interviews was chosen. Difficult for SMEs to keep the pace and stay competitive. Lack of skills to drive digital transformation, lack of strategic vision. Companies also have great difficulty funding new technologies and their rollout. As a result, the author recommends governmental action e.g., investments into digital infrastructure and into human capital by upskilling the workforce [6]. The research group around Anass et al. 2021 analyzed the relationship and application of the two concepts Lean and Industry 4.0 in the Moroccan context [7]. A survey was developed and conducted for this purpose focusing on the automotive, aerospace and chemicals sector. The researchers found that the concepts they analyzed are complementary to each other and can therefore be introduced simultaneously in companies. Furthermore, both concepts are considered important by industry.

The work of Faysse 2015 analysed the agricultural policy Green Morocco Plan, initiated in 2008 by the Moroccan Department of Agriculture and Marine Fisheries, based on policy document reviews and implemented instruments [8]. The goal of the Green Morocco Plan was to ensure economic growth within the next 10 - 15 years in the agriculture sector [8]. According to the author, the policy only succeeded moderately in developing and transforming rural areas to increase the agricultural production and thus also the individual income in these areas. The opportunities and challenges of digitalization in the Moroccan agriculture sector were discusses by Jabir & Falih 2020 [9]. For this purpose, latest technologies were identified and their applicability around the region Beni-Mellal Khenifra analyzed, e.g., Wireless Sensor Networks applied in different climates, energy harvesting technology, decision support software. To ensure success in the digital transformation and boost economic growth in agriculture and food industry the authors recommend good governance and policy.

3. Competence Center on Automation (CCoA)

The CCoA is intended to serve as a platform for transferring knowledge in the field of automation and digitalization in an applied manner using new approaches to action-oriented learning. For example, to realize the potential of university-industry-collaborations, a new educational strategy is needed that combines existing concepts and didactic approaches [10]. For the further training of specialists and executives, real-world projects can as well be used according to inquiry-based learning approaches [11].

The focus of the CCoA is during its pilot phase laid on two key industries, automotive and food, to develop differentiated and demand-oriented solutions. The automotive industry plays an important economic role not only for Morocco, but for the entire North African region [12]. Therefore, investments are made not only by the individual states and companies but also by the Africa Bank of development, European Union, World Bank and others. In addition to the high efforts of the Moroccan government by establishing several free trade zones such as Tangier Automotive City and Kenitra Atlantic Free Zone, there has also been a high level of investment from abroad [13]. In the last 5 years, many investments have been made by foreign automotive companies, including French, Spanish, Japanese and US corporations. On the one hand, the Moroccan market is used as a cost-effective production location to supply automotive systems and -parts to plants spread around the world. On the other hand, the African and especially the Moroccan sales market is gaining more and more interest.

Furthermore, the food sector including agriculture and beverage is economically very important and generates revenues over USD 12,8 billion [14]. According to FENAGRI [14], Morocco's export in the food sector is increasing especially for granulated sugar, fresh fruits, and fresh tomatoes.

Industries such as consumer electronics and energy are also gaining high interest in Morocco but are not focused in this pilot phase of the CCoA.

Innovative and comprehensive education and training seminars are planned within the project scope. Hence, seminars for experienced professionals and managers are planned, as well as trainings for students and graduates in learning factories [15]. Further, industrial train-the-trainer seminars for system integrators are also scheduled.

4. Methodology

Quantitative and qualitative elements are applied to identify the status quo and the need for automation and digitalization in Morocco. For the quantitative analysis, a comprehensive online questionnaire was developed. To take local conditions and requirements into account, the work was carried out by an interdisciplinary team of experts from Morocco and Germany. In addition to expertise in the engineering, natural- and social sciences, domain knowledge was contributed by local experts from Tangier in the automotive industries and from Casablanca in the food sector. The questionnaire was developed by the team in several workshops and review loops with industry experts. The online survey was set up and conducted using *QuestionPro*, a web-based software for market research and experience management. Current guidelines for the design, development and implementation of online surveys [16] have been taken into account. To ensure that as many participants as possible took part, the study was conducted bilingually in English and French. After the successful three-day test phase, the online survey started on 8th February 2021 and ended on 8th March 2021. To obtain a comprehensive picture in Morocco, three target groups were addressed, industry, universities, and vocational schools. Over all three online surveys, a cumulative total of 101 questionnaires were completed in full. The average completion time was 13 minutes. The dropout rate across all three surveys was 43,48 %.

In the following sections, the focus is on the industry questionnaire. The scope of this questionnaire included 46 questions. The limitations of this publication do not allow a detailed presentation of the entire survey; hence several interesting questions are discussed in detail here. Further information on the survey can be found on https://competence-automation.ma/.

5. Results & Discussion

The survey participants include experienced managers from industry networks of various local industry associations, personal contacts of local experts and the network of the German AHK (Foreign Chamber of Commerce) in Morocco. Figure 1 shows the industry distribution of the survey participants. The automotive sector accounts for the largest share of participants with 49 %. This high proportion in the automotive sector is partly due to the strong commitment of the industry association to the CCoA project. The automotive sector includes OEMs, system- and parts suppliers as well as service providers. Furthermore, with 20 % participation, the food sector represents the second largest group of participants. Within the CCoA project, agriculture and beverage are also subsumed under the food industry. The other industrial sectors individually do not account for more than 6%, as can be seen in Figure 1.

In addition to the industry distribution, the participant's function in the organization was also asked. More than 80 % of the participants hold management or executive positions. The revenue distribution of the participating companies presents a differentiated image. 25.9 % of the companies have revenues of $1 million or less per year. 20 small businesses with between 1 and 10 employees participated and represent about one-third of the industry participants. 36.5 % of the companies have revenues between $1 million and $50 million per year. 44 participating companies have more than 250 employees. A detailed analysis of the survey revealed that more than half of the participants were from international companies or Moroccan companies with foreign sites. 44 % of the participating companies are located only in Morocco. The survey therefore includes both SMEs and large companies.

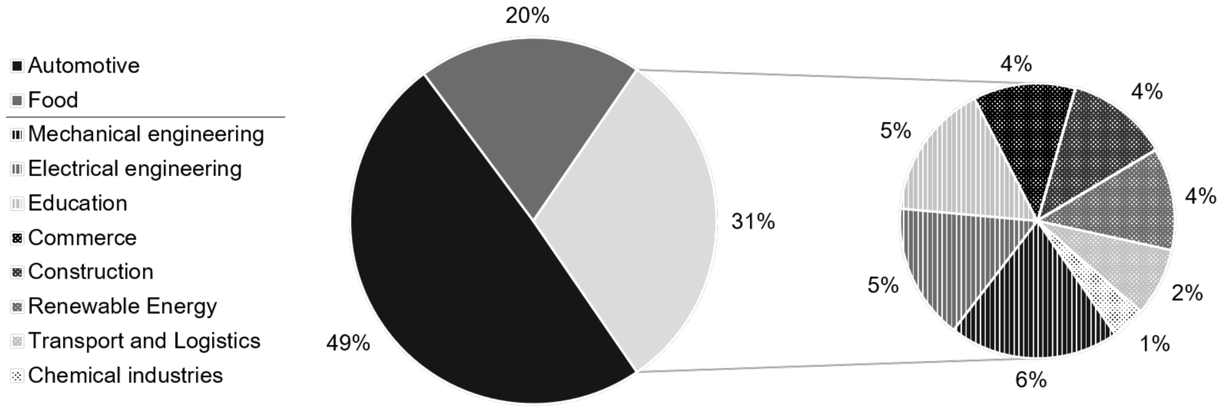


Figure 1: Industry distribution of the survey participants

In order to determine the status quo in the automation and digitalization field in Morocco, the As-Is Analysis team considered it very appropriate to ask the participants about the impact of global challenges on their organization. The following question was therefore asked for the industry's self-assessment:

- How would you rate the impact of the following challenges on your organization?

The results to this question are shown in Figure 2. 49 participants of the surveyed companies considered the impact of economic growth high till very high. Connectivity and Knowledge culture are also seen as important topics in the future, as 48 respectively 43 participants think it will have a high till very high impact. After analyzing the raw data, unexpectedly, the pollution challenge for the automotive industry is not as great in direct comparison to the food industry. This may be due to the less complex political conditions in the free trade zones. However, this must be investigated in further interviews with experts. Looking at the answers in more detail, the shortage of specialists is a greater challenge for the food industry than for the automotive industry. This will have a variety of reasons; from the point of view of local experts, the automotive industry finds it easier to attract skilled workers, as these often offer higher salaries, an international environment and thus a higher level of attractiveness.

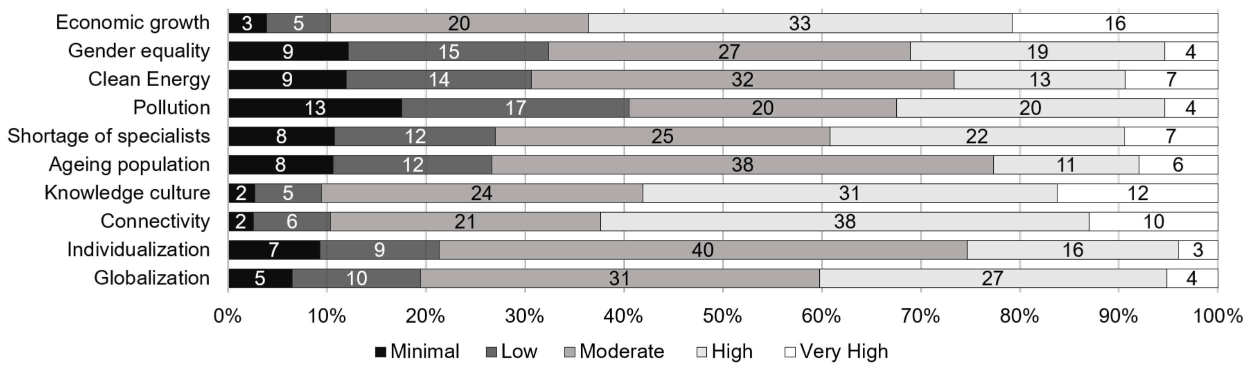


Figure 2: Impact of global challenges on the participating organizations

Furthermore, the following self-assessment question was asked of the survey participants:

- What is your experience with the following technologies?

Most of the surveyed participants are well experienced in electrical machines, industrial networks as well as network and communication. Many participants regularly use these technologies or even develop it further (see Figure 3). On the other hand, technologies like Blockchain, Process Mining, IoT (Internet of Things), RFID (Radio-Frequency Identification) and PLC (Programmable Logic Controller) are for half of the participants less known. Respectively they have not heard of these technologies or have heard of them but do not know more about them. Based on the self-assessment and detailed analysis of the raw data, it can be concluded that the food sector has less experience with the latest technologies.

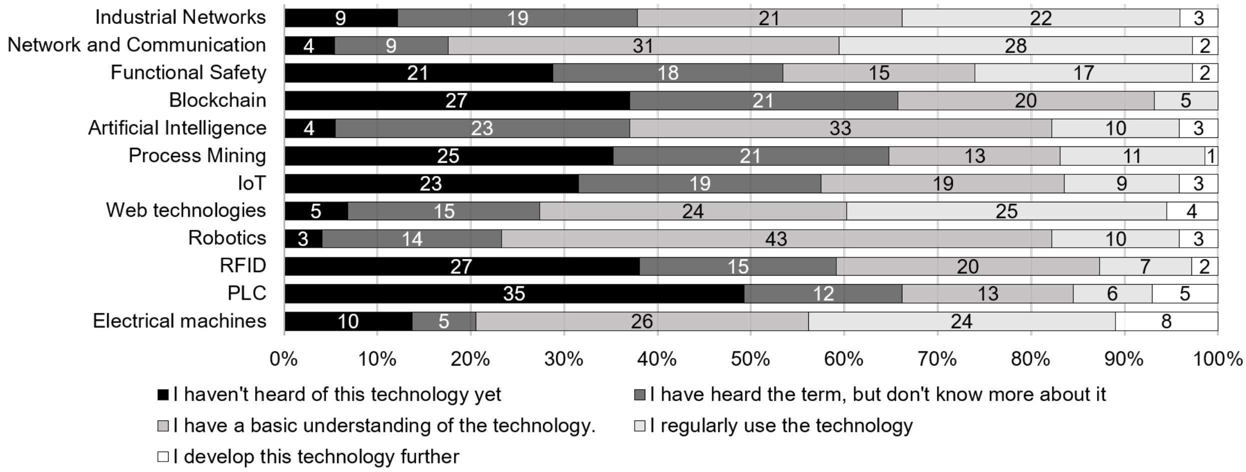


Figure 3: Self-assessment on experience with key technologies

Another key question on the status quo is covering the order management process:

- Is your order management process fully digital?

Cumulatively across all sectors, around half of the companies have already eliminated the need for paper orders and are handling them digitally. A closer look at the raw data reveals a large gap between the two industries, automotive and food, as shown in Figure 4. Over 63 % of companies in the automotive sector have implemented digital order management. However, only 31 % of companies in the food sector have implemented such a system.

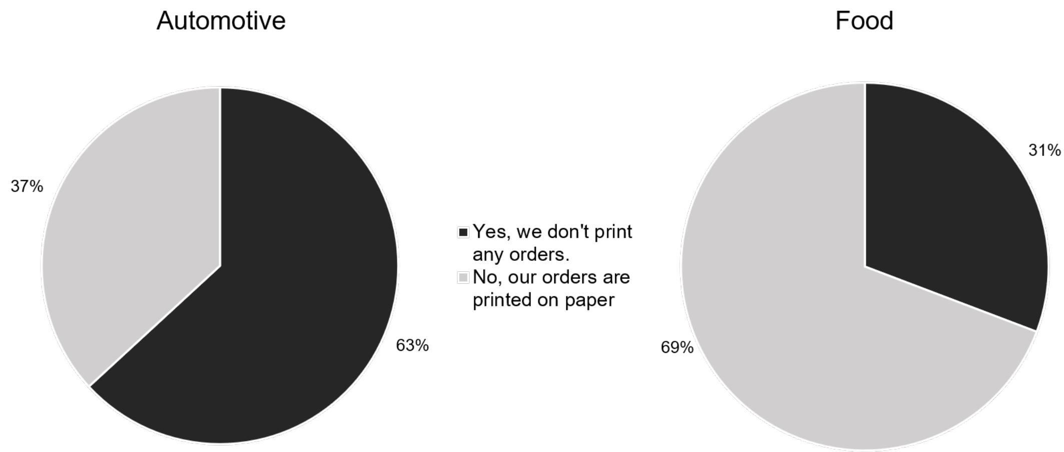


Figure 4: Order management paper based vs. digital.

Furthermore, questions were asked about the automation of quality control tasks. The automotive industry is fulfilling 28 % of quality control tasks manual, 69 % semi-automated and 3 % fully automated. In contrast, the food industry in Morocco is fulfilling 57 % of quality control tasks manual and just 43 % semi-automated. Once again, the automotive industry has made technological progress compared to the food industry and is already using fully automated quality control solutions in addition to semi-automated ones.

The use of professional Enterprise Resource Planning (ERP) systems is also dominated by the automotive industry (over 83 %) compared to the food industry (32 %). This is partly due to the fact that the automotive sector has a significantly higher share of multinationals in Morocco. These corporations often roll out this type of information and communication systems in a centrally driven manner. Thus, the decision and often the financing of these software solutions is made from industrialized nations.

The area of in-house training and continuing education of employees also shows a higher use of digital solutions, such as the use of E-learning platforms, by multinationals. The majority of international companies offer their employees E-Learning platforms. In contrast, only one-third of Moroccan companies offer E-Learning platforms for their staff. Again, Moroccan companies seem to be lacking in the handling and demonstration of the latest educational technologies as well as the financing of these.

6. Summary & Outlook

This publication provides an insight into the status quo in the field of automation and digitalization in the Moroccan automotive and food industry. Furthermore, the procedure of the quantitative and qualitative analysis was explained first. The results illustrate that, in contrast to the food industry, the maturity level of the automotive industry differs significantly with regard to automation and digitalization solutions. This is exemplified by the use of digital solutions in order processing or the use of ERP systems. Furthermore, figures showed that there is also a larger gap between Moroccan and international companies in terms of the use of new digitalization technologies but also the training of these. However, in general, there is a high level of interest in automation and digitalization solutions in Morocco. This was shown by the high level of participation in the survey and the subsequent feedback.

In a first step, the German-Moroccan commitment supported by the BMZ has already led to a higher sensitivity of the industry for automation and digitalization solutions. Furthermore, a high demand for professional education and training on new technologies could be identified. In the long term, the Competence Center on Automation will make a major contribution in the field of education and training and serve as a focal point to encourage bilateral industrial cooperation.

Acknowledgement

The establishment of the CCoA was supported by the “Partnership for Employment and Support for Medium-sized Enterprises in Morocco (PPE.MA)" Program, in the framework of the Special Initiative “Training and Job Creation” launched by BMZ, the Federal Ministry for Economic Cooperation and Development.

This paper and the research behind it would not have been possible without the support of the German BMZ and GIZ.

References

[1] UNCTAD, 2020, World Investment Report 2020, International Production Beyond the Pandemic, United Nations Conference on Trade and Development, New York.

[2] WMF, 2019, The 2019 World Manufacturing Forum Report – Skills for the Future of Manufacturing, www.worldmanufacturingforum.org, Milano.

[3] T. Friedli, G. Lanza, G. Schuh, C. Reuter, T. Arndt, B. Fränken, R. Leutzner, M. Wenking, 2015, Industrie 4.0 – ein Beitrag zur Entwicklung von "Smart Networks", zwf – Zeitschrift für wirtschaftlichen Fabrikbetrieb 110, Carl Hanser, Verlag, München.

[4] J. be Isa, T. Zimmermann, P. Scherwitz, G. Reinhart, 2020, Tactical Production Planning for Customer Individual Products in Changeable Production Networks, 1st Conference on Production Systems and Logistics CPSL.

[5] T. Bauernhansl, 2014, Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma, In: Industrie 4.0 in Produktion, Automatisierung und Logistik, Hrsg. T. Bauernhansl, M. ten Hompel, B. Vogel-Heuser, Springer Viewag Wiesbaden.

[6] A. Ben Hadj Hassine, 2021, The challenges of digital transformation: Case of Tunisian companies. Quantitative Economics and Management Studies, Vol. 2, No. 5, 281-287.

[7] C. Anass, B. Amine, H. Ibtissam, I. Bouhaddou, S. Elfezazi, 2021, Industry 4.0 and Lean Six Sigma: Results from a Pilot Study, In: A. Saka et al. (Eds.) Advances in Integrated Design and Production, LNME, pp. 613–619, Springer, Switzerland.

[8] N. Faysse, 2015, The rationale of the Green Morocco Plan: missing links between goals and implementation, The Journal of North African Studies, 20:4, 622-634.

[9] B. Jabir, N. Falih, 2020, Digital agriculture in Morocco, opportunities and challenges, IEEE 6th International Conference on Optimization and Applications (ICOA), Beni Mellal.

[10] A. Djalali, 2017, A didactic perspective on leadership education - focussing on the development of competencies within MBA programs. Steinbeis-Edition.

[11] W. Faix, L. Windisch, S. Kisgen, L. Paradowski, F. Unger, W. Bergmann, R. Tippelt, 2020, A new model for state-of-the-art leadership education with performance as a driving factor for future viability, In: Leadership, Education, Personality: An Interdisciplinary Journal, Vol. 2, pp. 59–74, Springer.

[12] T. Najar, M. Amami, 2012, A Multi-Case Study Approach of Interorganizational Information Systems Advantages in the Tunisian Automotive Components Supply Chain. International Business and Management, Vol. 4, No. 2, 92-110.

[13] AMICA, 2020, Industrie Automobile au Maroc, Tous pour un secteur Automobile marocain compétitif, attractif et visible à l'international, Presentation. http://amica.org.ma/. 09.04.2021.

[14] FENAGRI, 2017, Fédération Nationale de l'Agrolimentaire, Publications, Note commerce extérieur 2017, http://www.fenagri.org/Publication/1, 09.04.2021.

[15] L. Merkel, J. Atug, C. Berger, S. Braunreuther, G. Reinhart, 2018, Mass customization and paperless assembly in the learning factory for cyber-physical-production systems: Learning module 'from paperbased to paperless assembly', Proceedings - IEEE 18th International Conference on Advanced Learning Technologies, ICALT 2018, pp. 270–271.

[16] D. Andrews, B. Nonnecke, J. Preece, 2003, Conducting Research on the Internet: Online Survey Design, Development and Implementation Guidelines. International Journal of Human-Computer Interaction. 16, pp. 185–210.

Biography

Johannes be Isa (*1989) is an industrial engineer and expert in production networks. He is founder and general manager of apt advanced production training GmbH (apt).

Lukas Merkel (*1990) is a mechanical engineer and expert in IT systems in production systems. He is founder and general manager of apt.

Nicolas Brixle (*1996) is an industrial engineer and working as a project manager at apt.

John Ulrich Fimpel (*1958) is an industrial engineer and expert in international development cooperation. He is currently delegated by BMZ as business scout to the ZVEI, German Electric and Electronic Manufactures' Association.

Dr. Ardin Djalali, MBA (*1972) received his doctorate in the field of leadership education from the Ludwig-Maximilians-Universität München. He is currently Director International & MBA Programs at the Steinbeis School of International Business and Entrepreneurship (SIBE).

2nd Conference on Production Systems and Logistics

Determination Of The Level Of Automation For Additive Manufacturing Process Chains

Rainer Horstkotte[1], Benedikt Bruning[1], Marcel Prümmer[1], Kristian Arntz[1], Thomas Bergs[1,2]

[1]Fraunhofer Institue for Production Technology, Aachen, Germany
[2] Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Aachen, Germany

Abstract

Industrial manufacturing is confronted with increased cost pressure due to international competition. The use of automation solutions can help to optimally exploit existing potentials and react to market competitors. In particular, increased productivity and shorter cycle times lead to reduced costs and increased capabilities. New manufacturing technologies can also help to achieve an advantage over market competitors. In recent years, additive manufacturing technologies in particular have gained in importance.

Laser Powder Bed Fusion (L-PBF) is an additive manufacturing (AM) technology that enables the production of highly complex and individualized metal components. A significant disadvantage of L-PBF is the required post-processing of additive manufactured parts, which is necessary to remove auxiliary structures, separate the workpieces from the substrate plate and obtain high precision as well as low surface roughness. Automation of these post-processes is a crucial factor for increasing productivity and thus for further industrialization of L-PBF. In order to exploit this potential optimally, the level of automation has to be determined.

In this paper, a methodology is presented that enables the determination of the level of automation for the additive process chain with L-BPF. The focus is on evaluating the level of automation of individual manufacturing technologies due to consideration of technology-specific requirements and characteristics. The scope of the analysis is not limited to technologies; handling processes are also taken into account. A differentiated evaluation of the level of automation is enabled by the definition of technology-specific and cross-technology sub-tasks.

Keywords

Laser Powder Bed Fusion; L-PBF; automation; level of automation; process chain; additive manufacturing; AM; post-processing

1. Introduction

The increasing number of variants requires manufacturing to optimize flexibility. Addressing this demand, additive manufacturing processes are becoming more and more viable. The number of machines and printed parts sold has grown steadily in recent years due to the increased productivity and quality of AM machines and increasing market diffusion of the technology [1, 2]. In addition, manufacturing companies in high-wage countries such as Germany are facing increased competition due to the cost advantages of production in Asia and Eastern Europe. The application of automation solutions has the potential to significantly increase productivity and thus ensure competitiveness in the international market environment. The current state of the art in the industry shows that the automation potential in the additive process chain is not fully exploited. Especially in the area of post processing, the majority of process steps are currently conducted manually, which leads to high costs [3]. Since the costs incurred for the additive manufacturing of a component are compared to both conventional manufacturing methods and competing companies, the reduction of costs is an important priority for the economic use of additive manufacturing technologies in industrial manufacturing [3].

Determining the level of automation (LOA) is able to support the assessment of the status quo in terms of productivity. For improving the status quo, it is necessary to identify potentials and define fields of action. For this purpose, specific requirements for automation and special constraints of the additive process chain have to be considered. A methodology for the determination of the LOA enables the assessment of the current state and the identification of potentials for the significant improvement of productivity.

This paper presents a methodology that enables the determination of the LOA for the additive process chain with L-BPF. The focus is on evaluating the LOA of individual manufacturing technologies due to consideration of technology-specific requirements and characteristics. The scope of the analysis is not limited to technologies; handling processes are also taken into account. A differentiated evaluation of the LOA is enabled by the definition of technology-specific and cross-technology sub-tasks. In the second chapter, the initial situation concerning the additive process chain is presented using the example of the AM technology L-PBF and the need for a methodology to determine the current level and potential of automation is derived. The following third chapter describes the methodology that enables the determination of the LOA for the additive process chain. In the fourth chapter, the exemplary application of the methodology is briefly described. The final chapter five concludes the paper and presents possible directions for further research.

2. Initial Situation

With the number of additively built parts and products growing year after year, AM technologies are becoming increasingly important and shape the world of production in the future [4]. Based on the principle that a solid body is formed by joining material layer by layer using a chemical or physical process, additive manufactured components only have few minor design constraints such as the layers' thickness [2, 5, 6]. One of the most important and most used technologies in the group of AM technologies is L-PBF [1]. Here, the AM principle is realized by a layer-wise application of metal powder [1]. Molten powder is fused with the surrounding material by a laser, solidifies and thus forms the component layer wise [2, 5]. There are several advantages of L-PBF: Similar mechanical properties compared to parts conventionally manufactured, density of almost one hundred percent and new design possibilities for metal components and the integration of several functions into one single component allows individualization and topology or material optimization for lightweight constructions [7,8, 9, 10]. But there are also disadvantages of L-PBF that need to be reconsidered: For the additive manufacturing process itself, support structures and a baseplate are necessary to mitigate the warping and to dissipate heat [7]. Furthermore, the removal of unfused excess powder is necessary after the build process due to the risk of powder cross-contamination and associated quality degradation or adverse effects on the workers' health [4, 5]. A rough surface, significant shape deviations and residual stress can lead to warping or anisotropic mechanical properties [8, 11]. Therefore, L-PBF is used in combination with other conventional manufacturing technologies to ensure dimensional and positional accuracy of the manufactured components [5, 10, 11, 12].

The typical process chain of an additive manufactured component with its post-processes is described in the following. First, the components are built on a baseplate to which they are fused [7]. To form components, various powder layers are locally fused. The additive components are surrounded by unfused powder, which can be recycled also like the substrate plate [6]. The remaining powder must be removed. This is achieved by vacuuming and additional cleaning, but it has to be done within a glove box to safely handle the powder [7]. Recent powder removal-methods also consider ultrasonic cleaning equipment or kinematic powder removal systems [4]. To reduce stress resulting from the AM process, stress relief heat treatment is an option, but there exist also other possible heat treatments according to the requirements of the material [7, 10]. Afterwards, the component is separated from the baseplate by sawing or wire-EDM [11]. Additionally, auxiliary structures need to be removed by manual chiseling and grinding to smooth the surface [11]. Alternatively, milling or other automated subtractive processes can be used [13]. Furthermore, functional features are conducted, e. g. by milling, turning, sink- and wire-EDM [11]. To improve the surface quality, sand blasting or barrel finishing or other subtractive processes according to the component requirements are applied [11]. Apart from the technologies mentioned above, others may also be used [10].

The comparatively low productivity and high costs of such AM process chains are currently a barrier to the economic use of the technology [7]. Industrialization of L-PBF could be achieved by automating the L-PBF manufacturing process chain, as it reduces costs and increases productivity [14, 15]. For the development of automation concepts, an analysis of the status quo of the additive process chain is essential. Therefore, a discourse on scientific approaches to the topic of additive process chain and automation follows:

MÖHRLE considers the L-PBF process chain mainly from an economic and organizational point of view and does not explicitly include automation [11]. BÖCK's approach supports the integration of non-conventional technologies into conventional process chains, but does not address L-PBF and the specific requirements of automation [16]. The approach of PRÜMMER also does not fully consider the automation of the additive process chain. Here, an evaluation system for automated manufacturing systems in toolmaking is presented based on typical technologies in toolmaking. L-PBF and other AM technologies are therefore not considered [18]. In SEIFERMANN's concept, too, automation alternatives are evaluated and selected mainly on the basis of lead time and additional other economic and technological factors. A determination of the degree of automation is not explicitly made here [17]. At WINDMARK, automation alternatives are evaluated and selected based on economic factors. L-PBF and the specific challenges are not addressed [21]. Although KOPF explicitly considers the additive process chain with L-PBF in detail, automation and thus the determination of degrees of automation are only marginally included [20]. Currently, there are two main scientific approaches that explicitly deal with the definition of level of automation. On the one hand FAVRE-BULLE, who considers the level of automation as the quotient of the automated functions of a system in relation to all functions [15]. On the other hand, FROHM defines technological LOA in his approach, but these are general and do not consider the specific requirements of additive manufacturing [24].

The scientific discourse shows that although the approaches are applicable to several technologies, technology-specific characteristics are not taken into account. To achieve improvements in practice, it is necessary to identify technology-specific potentials and to derive requirement-oriented fields of action. For this purpose, the requirements for automation and the special characteristics of the additive process chain must be considered. Ideally, a methodology for determining the LOA can be used to evaluate the current LOA and identify potentials in order to achieve an optimal LOA. The following chapter presents such a methodology.

3. Methodology

Since the methodology should take into account the special characteristics of the additive process chain and the technologies used in it, a technology-specific consideration of the LOA is appropriate. This means that a separate LOA is determined for each technology used in the process chain. This results in a differentiated picture of the process chain. In addition, to consider the technology used in each case, handling functions that are relevant in the respective process step are also considered. The definition of the handling functions is based on the VDI standard 2860 [21]. Since the handling functions often rely on the same peripherals, they can be used in part across technologies to determine the respective LOA.

3.1 Determination of sub-tasks

Although the separate consideration of the technologies can already lead to the consideration of certain technology-specific characteristics, such a consideration cannot do adequate justice to the desired level of detail. Therefore, the individual process steps are subdivided into further sub processes. This procedure is based on Hierarchical Task Analysis (HTA), in which individual tasks are subdivided into sub-tasks. New sub-tasks can be identified if they pursue their own (partial) goal. However, it is not specified how detailed the sub-tasks must be subdivided. Thus a focus can be put on sub-tasks, which are to be considered more detailed, while others are subdivided less strongly. [22]

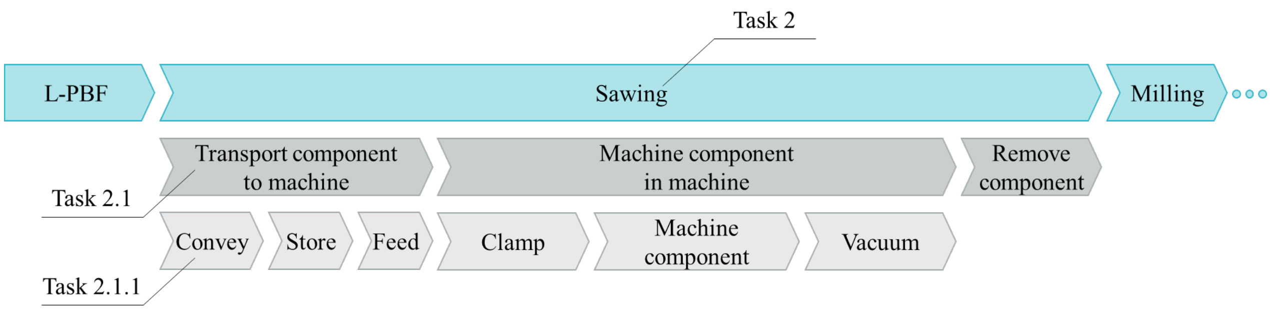


Figure 1: Hierarchical Task Analysis for sawing

Figure 1 shows an overview of how the HTA for sawing can look. The sawing process can be roughly divided into three steps within the process chain: The transport of the component to the machine, the machining of the component in the machine and the removal of the component. A consideration of the LOA at this level cannot yet be carried out in the desired level of detail, as it is not possible to clearly determine the extent to which these steps are automated or carried out manually. Therefore, a further detailing of the tasks is carried out. For the transport of the component to the machine, the sub-tasks conveying, storing and feeding, which are defined as handling functions, are suitable. In the task “Machine the component in machine”, all processes that take place in the immediate vicinity of the machine are taken into account. In order to be able to machine the component on the machine, it must first be clamped. The actual mechanical processing of the component represents a separate sub-task (“Machine component”). During the machining of the component, the extraction of chips can be used. For the removal of the component, only the removal is specified as a sub-task. The further transport of the component is not considered here, as it is considered as an upstream step in a subsequent technology. This ensures that identical processes along the process chain are not considered twice.

It is noticeable that not every task needs to be subdivided into sub-tasks at the same level of detail. While “Machine component in machine” is divided into four sub-tasks, “Remove component” isn’t further subdivided into sub-tasks. Tasks are only further subdivided if there are processes within a task that differ in terms of automation. It is not decisive whether the LOA of individual sub-tasks actually differs in the end, but merely whether a different LOA can result along the process chain. At this point, it should be noted whether different machines or peripheral systems are used between sub-tasks.

3.2 Weighting of sub-tasks

Since not every sub-task has the same significance for the LOA, the sub-tasks have to be weighted. For this purpose, target values are first defined that are important for automation. These are weighted afterwards with regard to the superordinate goal of the automation, whereby the sum of the goal is 100 %. In addition, the target values take into account the general conditions of the additive process chain. In this way, a weighting succeeds that is strongly adapted to the special characteristics of the additive process chain. Once the target values and the associated weights have been defined, the sub-tasks can be evaluated with regard to their influence on the respective target value. Based on the proficiency levels of the respective sub-tasks, a weighting factor $w_{x,j}$ can be determined. The calculation is shown in following formula:

$$w_{x,i} = \sum_{j=1}^{m} w_{j,y}\, b_{j,i} \tag{1}$$

The sum of the individual weights is normalized to 100 %. Finally, the weighting factors can be used to summarize the automation levels of the sub-tasks (cf. figure 2).

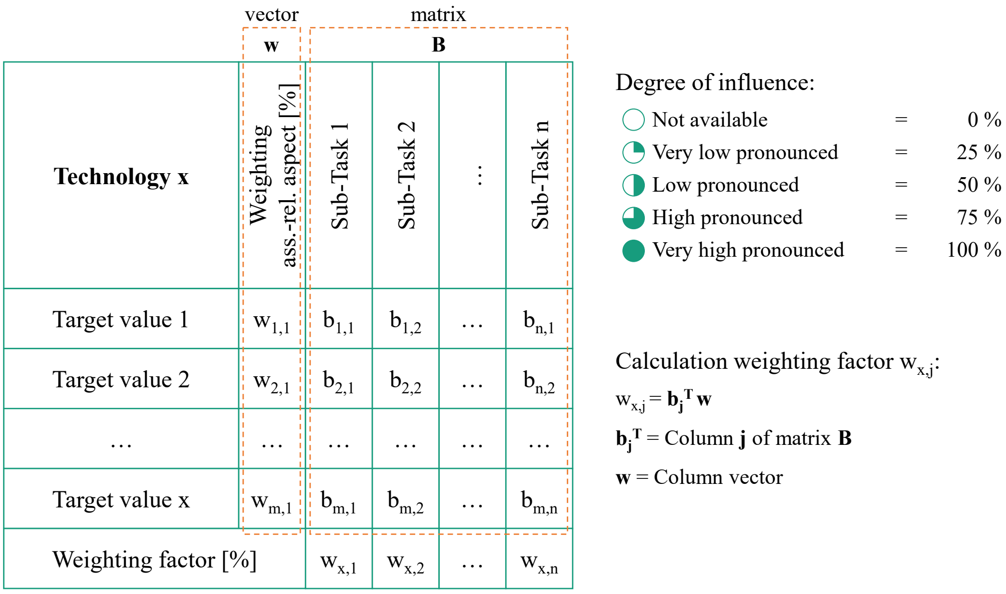


Figure 2: Determination of the weighting factors of the sub-tasks [23]

3.3 Definition of the scales for the determination of the automation level

The next step is the linkage of automation level to the individual sub-tasks. Therefore, sub-task-specific scales are developed in the methodology. In this way, it can be ensured that the special characteristics with regard to automation are taken into account for each sub-task. In this respect, the technology-specific approach with the possession of the LOA for each technology differs from approaches that only consider the entire process chain [7].

Following the Dynamo model, the scales for each sub-task are divided into two areas. One scale refers to the mechanical functions as well as the equipment used to conduct the sub-tasks, while the other scale refers to information processing and control [24]. The two scales are independent of each other. Thus, a differentiated view of the task division between employee and automation solution is achieved. To ensure a uniform size of the scales, a seven-level ordinal scale is developed for each sub-task, in which the highest LOA to be achieved is evaluated with the value seven and the lowest with the value one. The levels of each scale are determined based on possible LOAs. It may happen that not every one of the seven levels in the scale is occupied. In these cases, the increase in the LOA is not the same between each proficiency level, but this is taken into account when creating the scales.

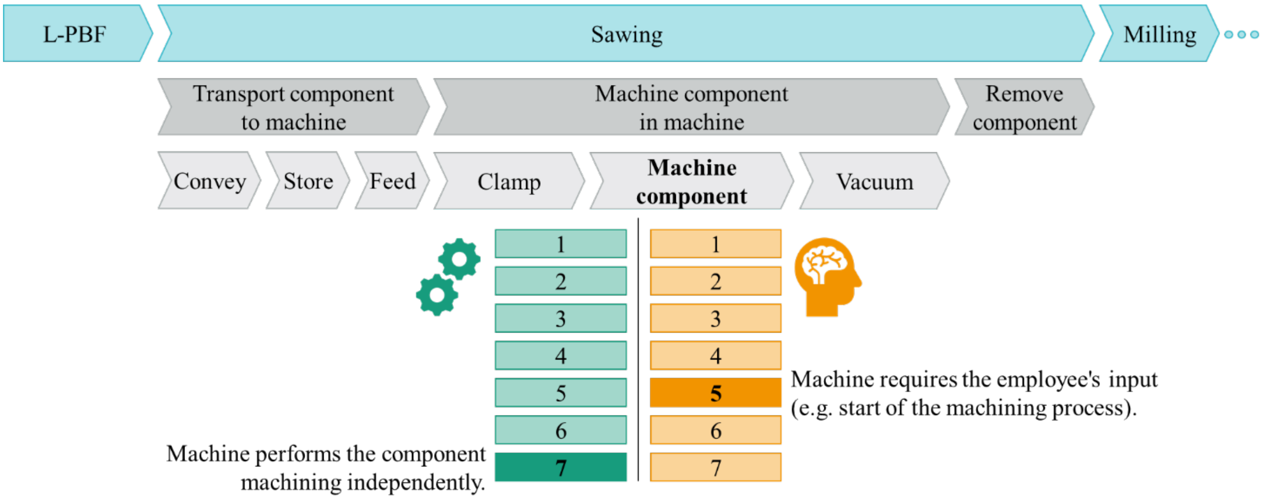


Figure 3: Exemplary representation of the automation scale of a sub-task for sawing

Figure 3 illustrates the automation scale of the sub-task "Machine component" from the process step "Sawing". The exact definition for assigning a particular value on the scale must be determined in advance for each subtask. Gradations of the scale for mechanics and equipment differ with regard to the tools used and the physical demands on the employees. For the scale for information and control, the decisive factor is what information is available to the employee or whether the work orders are clearly defined so that uniform execution can be ensured independently of the employees. For this, it must also be considered whether the employee makes independent decisions and thus has a strong influence on the execution of the sub-task. In the example, the machine performs the machining of the component independently. Therefore, with regard to the scale for mechanics and equipment, the highest value is assigned. Since the machine requires the employee's input, e. g. to start the machining process, for the scale of information processing and controls a lower value is chosen.

3.4 Determination of the overall degree of automation

After defining the sub-tasks with the associated weightings and the sub-task-specific scales for determining the LOA, a technology-specific automation level can be calculated. For this purpose, the automation levels of the sub-tasks are added according to their weighting. The designation LOA stands for the calculated LOA. The LOA is calculated for both the mechanical and the cognitive (information processing and control) level.

$$LOA_x = \sum_{j}^{n} w_{x,j} * LOA_{x,j} \qquad (2)$$

By determining the LOA of a technology, it is possible to compare the automation in different process chains in relation to the technology under consideration. A comparison between different technologies does only work to a limited extent, since each technology is based on different evaluation scales. However, the cross-technology comparison helps to determine the extent to which automation potentials have been exhausted. In order to identify imbalances in the degree of automation within a process chain, an overview is created that shows the different technologies in terms of their LOA.

In the example shown in Figure 4, all sub-tasks are first evaluated with the two scales. The automation levels of all sub-tasks are then added to form a technology-specific automation level. The overview shows how the LOA of individual technologies can differ from one another. The technologies do not necessarily have to be equipped with different automation technologies. The minima and maxima of a technology realizable in each case can also be decisive for the differences. The overview therefore serves rather to identify the extent to which the automation potentials have been exhausted.

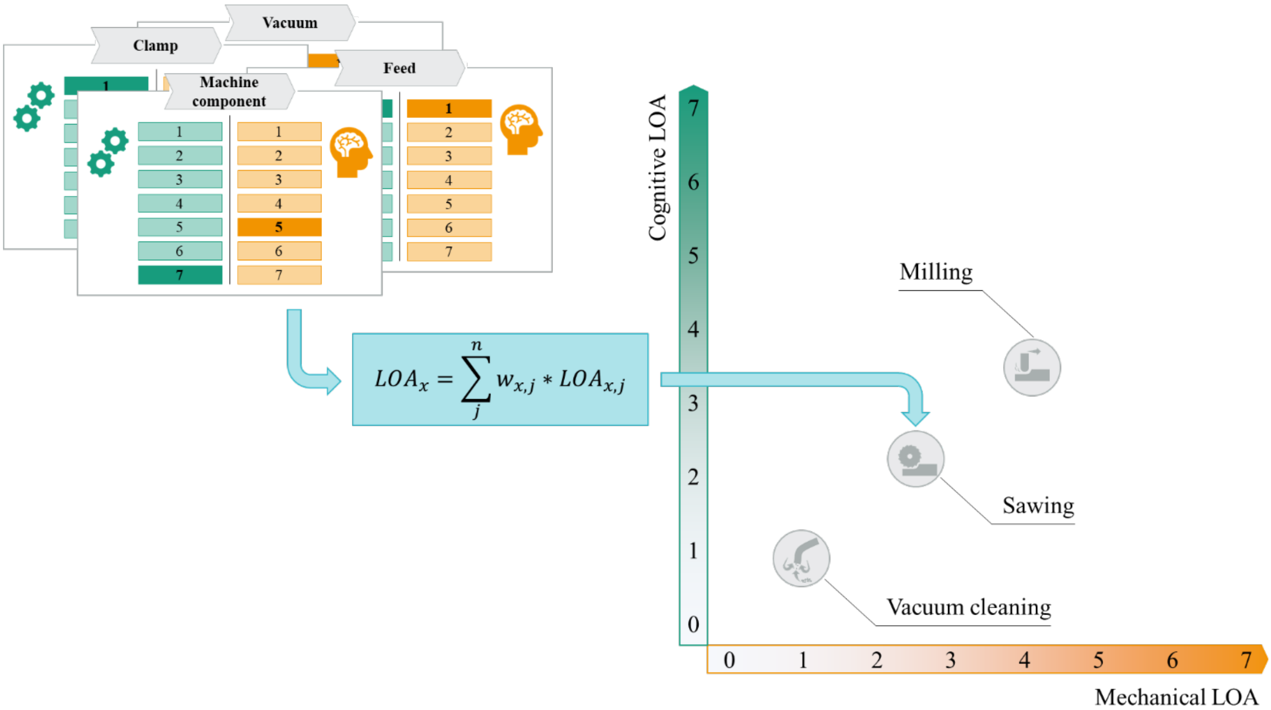


Figure 4: Overview of technology-specific automation levels

4. Application

The developed methodology was applied to an exemplary additive process chain. Each process step was divided into sub-tasks and the corresponding automation level was determined based on the respective scale (maximum of 7 points possible for the mechanical and cognitive classification). With the help of the determined weighting factor, a calculation was possible to determine the total LOA of a process step or technology. Figure 5 provides an overview of the LOA of all the technologies considered.

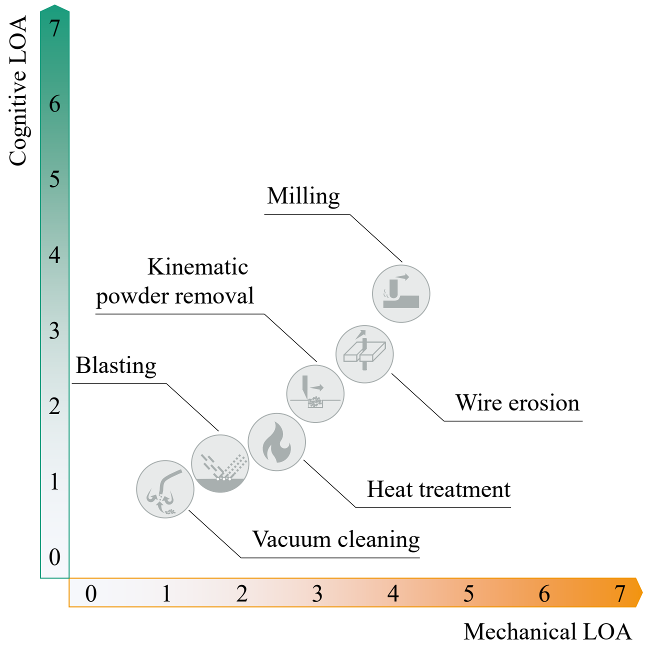


Figure 4: Overview of LOA of technologies considered in the application

Measured by the spectrum of the scale, it is noticeable that the LOA of the technologies considered is relatively low. A higher degree of automation can be seen in some technologies that rely on a static machine (wire erosion, milling). This can be attributed to the equipment of the machines used and leads to a partial automation of some sub-tasks. It is also noticeable that the mechanical LOA is higher than the cognitive degree for all the technologies considered. A stronger focus on the automation of mechanical functions may be one reason for this observation. When considering the LOA, however, it should be noted that the highest LOA cannot be achieved for every technology, since optional sub-tasks are not considered in some cases. Too great a difference in the automation levels of successive sub-tasks would not make sense from an economic point of view, since in this case unmanned production would not be possible despite the automation of individual sub-tasks. A detailed examination of the sub-tasks can also provide information about which higher levels of automation exist and thus serve as a basis to increase the LOA.

5. Conclusion and Outlook

At present, however, components produced using L-PBF cannot be manufactured to a satisfactory quality for immediate use. Therefore, further processing of the components in the additive process chain cannot be dispensed with. A challenge to the use of L-PBF is the high proportion of manual work steps in the further processing of the components, as this makes the entire process chain less economical in many cases. As a result, companies in high-wage countries such as Germany have difficulty withstanding the cost pressure from international competition. Automating the post processing of additively manufactured components can help to reduce the costs of the process chain.

In order to provide a basis for further decisions regarding automation, a methodology has been developed to determine the LOA of the additive process chain. This can be seen as an analysis of the current state and can be used in practice to enable further development of automation solutions. By using different models and methods, it was possible to develop a methodology for determining the degree of automation of an additive process chain. The four essential steps of the methodology for determining the degree of automation were presented in the paper. A distinction was also made between the development and the application of the methodology. Within the scales, the special characteristics of additive manufacturing are considered. In addition, the weightings are designed to match the objectives of automated additive process chain. Since the identification of sub-tasks considers not only the technologies , but also the associated handling functions, a comprehensive view of the value-adding and non-value-adding processes within the process chain can be ensured. Another advantage of the methodology is that it can be easily extended to include other technologies. Due to the large variety of technologies in the additive process chain, any extension may become relevant.

In this paper, the first approach for determining the LOA has been developed, which is explicitly designed for application in the additive process chain. The developed methodology represents a first step to develop an holistic automation strategy. Although measures for improving the LOA can be derived from the methodology, there is no monetary consideration of the automation solutions or no systematic approach. In order to increase the economic efficiency of the additive process chain, it may be useful to consider economic parameters. Moreover, the consideration of other technological parameters can also be beneficial, since maximizing the LOA is not the same as the optimum LOA.

Acknowledgements

Parts of the results were achieved in the project IDEA, support code 13N15001, funded by the Federal Ministry of Education and Research BMBF, Germany.

References

[1] Wohlers, T., 2019. Wohlers Report 2019: 3D Printing and Additive Manufacturing State of the Industry. Wohlers Associates, Inc, Fort Collins, Colorado.

[2] Dzemko, M., Engelmann, B., Hartmann, J., Schmitt, J., 2019. Toward Shifted Production Strategies Through Additive Manufacturing: A Technology and Market Review for Changing Value Chains. Procedia CIRP 86, p. 228.

[3] Zeyn, H., 2017. Industrialisierung der Additiven Fertigung. Digitalisierte Prozesskette - von der Entwicklung bis zum einsetzbaren Artikel. VDE-Verlag GmbH, Berlin.

[4] Wohlers, T., Campbell, R.I., 2017. Wohlers report 2017: 3D printing and additive manufacturing state of the industry annual worldwide progress report. Wohlers Associates, Inc, Fort Collins, Colo.

[5] Wollbrink, M., Prümmer, M., Kristian, A., Bergs, T., 2019. Methodology for Integration of Additive Manufacturing in Tool Manufacturing Process Chains. Proceedings of the 18th International Manufacturing Conference in China.

[6] Gibson, I., Rosen, D.W., Stucker, B., 2010. Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing. Springer Science+Business Media LLC, Boston, MA.

[7] Horstkotte, R., Heinrich, F., Prümmer, M., Arntz, K., Bergs, T., 2021. Generation and evaluation of automation concepts of additive process chains with Laser Powder Bed Fusion (L-PBF). Procedia CIRP 96, p. 97-102.

[8] Gebhardt, A., 2016. Additive Fertigungsverfahren: Additive Manufacturing und 3D-Drucken für Prototyping - Tooling - Produktion, 5th edn. Hanser, München.

[9] Klahn, C., Leutenecker, B., Meboldt, M., 2014. Design for Additive Manufacturing – Supporting the Substitution of Components in Series Products. Procedia CIRP 21, p. 138.

[10] Verein Deutscher Ingenieure. 3405-2 Additive Fertigungsverfahren. Strahlschmelzen metallischer Bauteile: Qualifizierung, Qualitätssicherung und Nachbearbeitung, Berlin. Beuth Verlag GmbH, 2013.

[11] Möhrle, M., 2018. Gestaltung von Fabrikstrukturen für die additive Fertigung. Springer Vieweg, Berlin.

[12] Wollbrink, M., Maslo, S., Zimmer, D., Abbas, K., Arntz, K., Bergs, T., 2020. Clamping and substrate plate system for continuous additive build-up and post-processing of metal parts. Procedia CIRP 93, p. 108–113.

[13] Denkena, B., Dittrich, M.-A., Henning, S., Lindecke, P., 2018. Investigations on a standardized process chain and support structure related rework procedures of SLM manufactured components. Procedia Manufacturing 18, p. 50

[14] Möhrle, M., Müller, J., Emmelmann, C., 2017. Industrialisierungsstudie Additive Fertigung: Herausforderungen und Ansätze. RTeJournal - Fachforum für Rapid Technologie.

[15] Favre-Bulle, B., 2004. Automatisierung komplexer Industrieprozesse: Systeme, Verfahren und Informationsmanagement. Springer Vienna, Vienna, s.l

[16] Böck, J., 2013. Methode zur Integration nicht-konventioneller Verfahren in flexible Fertigungs- und Montagelinien. Fraunhofer Verlag, Stuttgart

[17] Seifermann, S., 2018. Methode zur angepassten Erhöhung des Automatisierungsgrades hybrider, schlanker Fertigungszellen, 1st edn. Shaker, Herzogenrath.

[18] Prümmer, M., 2020. Kennzahlenbasiertes Bewertungssystem der Leistungsfähigkeit verketteter Fertigungssysteme: In der mechanischen Fertigung des Werkzeugbaus. Apprimus Wissenschaftsverlag, Aachen.

[19] Windmark, C., 2018. Performance-based costing as decision support for development of discrete part production: Linking performance, production costs, and sustainability, Lund.

[20] Kopf, R., 2018. Kostenorientierte Planung von Fertigungsfolgen additiver Technologien, 1st edn.

[21] Richtlinie VDI 2860, 1990. Montage- und Handhabungstechnik.

[22] Kirwan, B.; Ainsworth, L. K., 1992. Guide To Task Analysis. The Task Analysis Working Group. CRC Press, London.

[23] Bilsing, A., 2006. Kennzahlengestützte Bewertung der technologischen Leistungsfähigkeit der Fertigung im Werkzeug- und Formenbau. Diss. RWTH Aachen, Aachen.

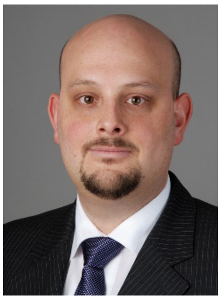
[24] Frohm, J.: Levels of automation in production systems. Zugl.: Göteborg, Univ., Diss., 2008. Göteborg: Chalmers Univ. of Technology, 2008

Biography

Rainer Horstkotte (*1992) has been research fellow at Fraunhofer-Institute for Production Technology IPT since 2018. In addition, he is head of the business unit tool and die making. His specific area of research is the automation of additive process chains.

Benedikt Bruning (*1995) has been scientific assistant at Fraunhofer-Institute for Production Technology IPT since 2017. He finished his studies in mechanical engineering in 2021. His specific area of research is the automation in manufacturing and the determining of level of automation.

Marcel Prümmer (*1989) has been head of the Research Group Technology Organization at Fraunhofer-Institute for Production Technology IPT since 2018. In addition, Dr.-Ing. Marcel Prümmer is division manager at the WZL Aachen GbmH for the area of manufacturing. His specific area of research is the automation in manufacturing.



Kristian Arntz (*1970) has been head of department Technology Planning at Fraunhofer-Institute for Production Technology IPT since 2009. In addition, Dr.-Ing. Kristian Arntz is manager R&D at WBA Aachener Werkzeugbau Akademie GmbH. His specific area of research is milling in tool and die making and technology planning.

Thomas Bergs (*1967) has been director of the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since 2018. In addition Prof. Dr.-Ing. Thomas Bergs, MBA is director of the Fraunhofer-Institute for Production Technology IPT. His specific area of research is the manufacturing technology.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Analysis of surface quality during milling with industrial robots as a function of milling spindle orientation

Lars Niklas Penczek[1], Michael Krampe[1], Bernd Kuhlenkötter[1]
[1] Ruhr University Bochum, Chair of Production Systems, Bochum, Germany

Abstract

The range of tasks performed by industrial robots has become increasingly comprehensive in recent decades. This includes pick-and-place tasks with low to medium handling weights, as well as material processing tasks with medium to high processing forces. Milling in particular involves high-frequency excitation of the material, which results in vibration phenomena and therefor affects the surface quality of the product. These effects were investigated by milling tests at characteristic positions and in different orientations of the milling spindle, using aluminum. For this purpose, based on an experimental modal analysis, a prediction model of the robot's natural frequencies in dependence of the joint positions was developed. This allowed the determination of characteristic milling positions and milling spindle orientations. Subsequently, the milling results were related to a causal connection and were analyzed regarding the natural frequencies.

Keywords

Industrial Robotics; Milling; Tool Orientation; Vibration Analysis; Experimental Modal Analysis; Ball-End Milling Cutter; Regression Model

1. Introduction

The application areas for industrial robots now extend far beyond the classic welding, assembly and bonding tasks. For example, industrial robots are increasingly being used in metalworking applications, such as polishing [1], sheet metal forming [2], deburring, cutting, drilling, grinding and milling. In contrast to CNC machines, articulated robots offer high flexibility, low investment cost per cubic meter of workspace, as well as possible interaction with other machines or additional axes [3]. The machining forces in robots are in a medium to high range, which can lead to a high displacement of the end effector [4] as well as vibration phenomena [5], occurring due to the up to 50 times lower stiffness compared to CNC machines [6]. Especially the milling of hard materials such as metals or ceramics therefor is a great challenge.

The inaccuracies in milling with industrial robots originate in the mechanics, the drives, the programming, the workpiece, the process and the tool. The geometric errors (zero position error, arm length and angle errors) have the greatest influence on the dimensional accuracy and the surface quality of the milled components [7]. Especially the surface quality of the workpieces depends on the vibration phenomena resulting from the excitation of the system and the natural frequencies and modes. Redundant robot poses can be used to influence this effect. The use of ball-end milling cutters makes it possible to extend this redundancy enormously, since the orientation of the milling spindle can be adjusted in all three directions of rotation. This potential will be systematically investigated for the first time in this paper, using a prediction model to determine the natural frequencies. The influence of self-excited vibrations on the milling pattern was investigated in milling tests during finishing.

2. Milling with industrial robots

Nowadays, industrial robots are used in industrial applications and in the cultural industry for the machining of plastic, hard foam, wood and occasionally aluminum parts, where primarily parts with medium quality are milled. Especially for demanding parts with complicated free-form surfaces, jointed-arm robots offer significant advantages in terms of accessibility. The process of milling is basically divided into the machining steps roughing and finishing. During roughing, the chip removal is very large, which means that the process forces that occur and the displacement of the tool center point (TCP) are also particularly high. In finishing, the component is finely machined and high-quality surfaces with narrow tolerances are milled [8], so that any possible surface defects caused by vibrations, are immediately visible in the finished workpiece. In form milling, a spherical cutter is often used in the finishing process, which has a variable approach of the cutter to the surface and thus allows a high number of possible joint configurations.

Robots are particularly susceptible to vibration, because the stiffness for a six-axis industrial robot is typically less than 1 N/μm, while standard CNC machines often have a stiffness of more than 50 N/μm [6]. The vibrations occurring during the milling process can be divided into self-excited and externally excited vibrations. Externally excited vibrations are caused by a multi-bladed tool that is not continuously engaged and thus exerts a periodic force on the robot [9]. In this case the vibration frequency is determined by the excitation frequency. If this frequency is in close proximity to the natural frequencies of the robot, which are usually in the range of about 10 Hz [6], an upswing of the system can be triggered. Therefore, in order to describe the vibration effects comprehensively, a consideration of the excitation frequencies is also useful and will be investigated in more detail in future work.

Among the self-excited oscillations, those resulting from the dynamic process itself and the energy of the robot system's drives, are particularly problematic. Even if no natural frequencies of the system are directly excited by the periodic application of force, the robot still oscillates in its natural frequencies. This leads to wavy contours on the surface of the material, so-called chatter marks. The undulating surface further amplifies this effect for each chip, resulting in self-exciting oscillations.

Both types of vibrations mentioned above are summarized as chatter vibrations in literature. Chatter vibrations do not only reduce the quality of the workpiece, they also drastically increase the load on the tool and milling robot. In unfavorable cases, this can result in damage to the workpiece, tool or robot [10].

3. Model for natural frequency determination in a defined workspace

In order to understand the mechanical oscillation of a robot during milling, the corresponding influencing variables must be determined over the entire workspace. These influencing variables are the natural frequencies of the robot, the excitation frequencies and the milling forces that occur. The natural frequencies vary, depending on the joint position of the robot. In order to determine these natural frequencies, an experimental modal analysis was carried out at characteristic points and a regression model was then created to extend the results over the entire workspace.

Experimental modal analysis (EMA), is the most common method for characterizing the dynamic behavior of mechanical structures. It is built on the foundation of calculating the transfer function $G(\omega)$, which represents the sought linear relationship between the input and output signals of a system in the frequency domain. This function is calculated from the quotient of the excitation spectrum $U(\omega)$ and the resulting system response $Y(\omega)$. The applied excitation force leads to a structural response, which is measured as acceleration amplitude. This excitation force can be triggered continuously by an electrodynamic excite or impulsively by a hammer blow with force pulse measurement [11].

In order to investigate the natural frequencies, a statistical design of experiments was used. In these experiments specific robot configurations were investigated [12,5], which in the context of this modal analysis, were limited to a working range that can be used for the milling process. A centrally composed experimental design was applied, based on a fully factorial experimental design [13]. The nonlinear relationships between the natural frequencies were investigated by adding star and center points. The orthogonal, centrally composed experimental design now results in a factor level combination table for four factors with three factor levels each. Thus, the number of measurements to be investigated could be limited to a total of 27 joint configurations.

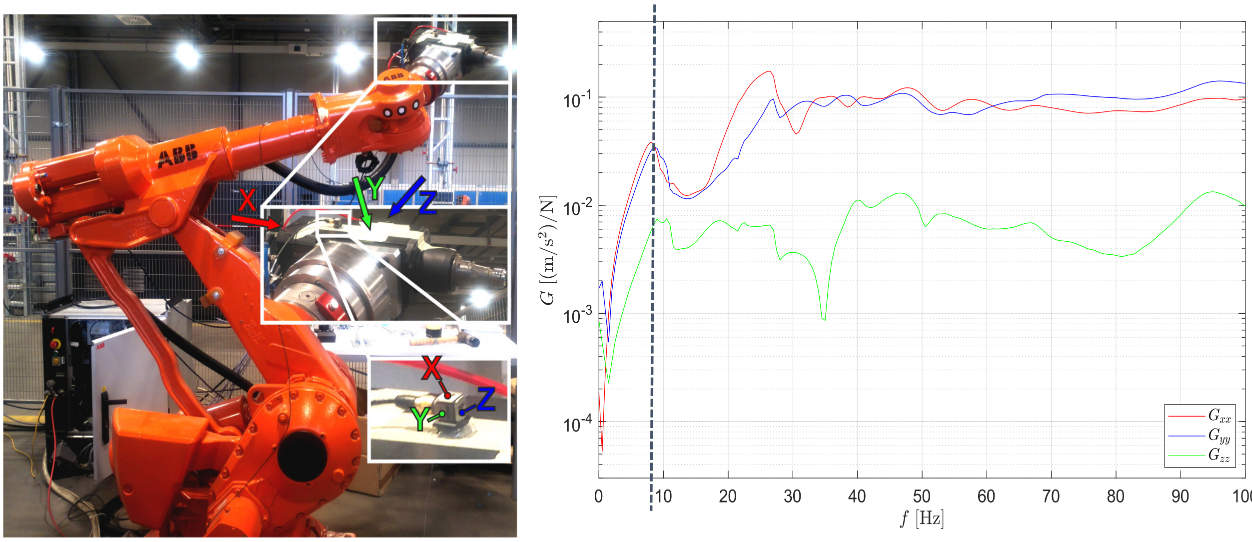


Figure 1: Joint position of the ABB IRB 4400 (left; with force excitation direction) and transmission functions G_{xx}, G_{yy} and G_{zz} as a result of the EMA (right) with the joint configuration: $p_1 = 10, p_2 = 40, p_3 = -15,\ p_4 = -50,\ p_5 = -45,\ p_6 = 0$

The modal analysis was carried out on the *ABB IRB 4400* six-axis robot, which has additional stiffening of the second link in its design, resulting in a particularly positive effect on milling tasks. The heavy-duty impulse hammer Type 8208 from *Brüel & Kjaer* was used to excite the robot, which generates an excitation force of up to 44.4 kN and guarantees a resolution of 0.225 mV/N. The triaxial accelerometer Type 4524-B-001 from *Brüel & Kjaer*, possessing a sensitivity of 10 mV/g, was used to record the resulting systemic response. The system was excited at one measuring point in each of the three coordinate directions, using the single impact method directly on the milling spindle, see Figure 1. The triaxial acceleration sensor was also placed close to the milling spindle, in order to be able to measure the frequencies as close as possible to the TCP.

As already mentioned, a natural frequency in the range of approximately 10 Hz is to be expected, therefore the recording range was set to 0 - 100 Hz. The transfer function was calculated by averaging the measurement results from 5 measurement processes. These measurements showed that the natural frequencies of the robot ranged from 8.4 Hz to 17.9 Hz.

In order to transfer the natural frequencies to the entire workspace and to all poses of the robot, a regression model was developed. From previous investigations in six-axis robots [14] it is known that joints 1 and 6 only have a very small influence on the natural frequency of the robot, compared to other joints. Therefore, they can be neglected for further consideration, so that only the effects of joints 2, 3, 4 and 5 were considered in the model design. The natural frequencies of the robot exhibit a strong nonlinear dependence on the joint configurations [13], therefore a quadratic model was used to empirically describe the relationship between the target variable y_i and the factors x_1, x_2, x_3 and x_4. For this purpose, the target variable is described by the following term:

$$\vec{y} = \underline{X} \cdot \vec{\beta} \tag{1}$$

where $\vec{y} = [y_1 \quad y_2 \quad \cdots \quad y_n]^T$ as well as $\vec{\beta} = [\beta_{44} \quad \beta_{34} \quad \beta_{33} \quad \beta_{24} \quad \cdots \quad \beta_0]^T$.

Also applies

$$\underline{X} = \begin{bmatrix} x_{41}^2 & x_{41} \cdot x_{31} & \cdots & 1 \\ x_{42}^2 & x_{42} \cdot x_{32} & \cdots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{4n}^2 & x_{4n} \cdot x_{3n} & \cdots & 1 \end{bmatrix}. \tag{2}$$

Therefor

$$y_i = f\left(x_{4i}, x_{3i}, x_{2i}, x_{1i}, \vec{\beta}\right) = [x_{4i}^2 \quad x_{4i} \cdot x_{3i} \quad \cdots \quad 1] \cdot \vec{\beta}. \tag{3}$$

Here $\vec{x_i} = [x_{1i} \quad x_{2i} \quad x_{3i} \quad x_{4i}]^T$ denotes the joint variables $[q_2 \quad q_3 \quad q_4 \quad q_5]^T$ from the i-th joint configuration and y_i corresponds to the associated natural frequency. The function $f\left(\vec{x_i}, \vec{\beta}\right)$ now depends on the parameter vector $\vec{\beta}$ which can be estimated by using the method of the smallest squares [15]. In the method of the smallest squares, the parameter vector $\vec{\beta}$ is defined to minimize the sum of squared errors. This results in

$$\min_{\vec{\beta}} J\left[f\left(\vec{x}, \vec{\beta}\right)\right] = \sum_{i=1}^{n} \left[y_i - f\left(\vec{x_i}, \vec{\beta}\right)\right]^2 \tag{4}$$

If the quality functional $J\left[f\left(\vec{x}, \vec{\beta}\right)\right]$ is derived according to parameter vector $\vec{\beta}$. The result is

$$\frac{dJ}{d\vec{\beta}} = -2\underline{X}^T \vec{y} + 2\underline{X}^T \underline{X} \vec{\beta} \tag{5}$$

For the local and global minimum of the quality functional applies $\frac{dJ}{d\vec{\beta}} = 0$.

This results in [15]

$$\vec{\beta} = (\underline{X}^T \underline{X})^{-1} \underline{X}^T \cdot \vec{y}. \tag{6}$$

In this calculation, the Hessian matrix $\frac{d^2 J}{d\vec{\beta}^2} = 2\underline{X}^T \underline{X}$ is always positively defined, providing the optimal result for the calculated parameter vector $\vec{\beta}$.

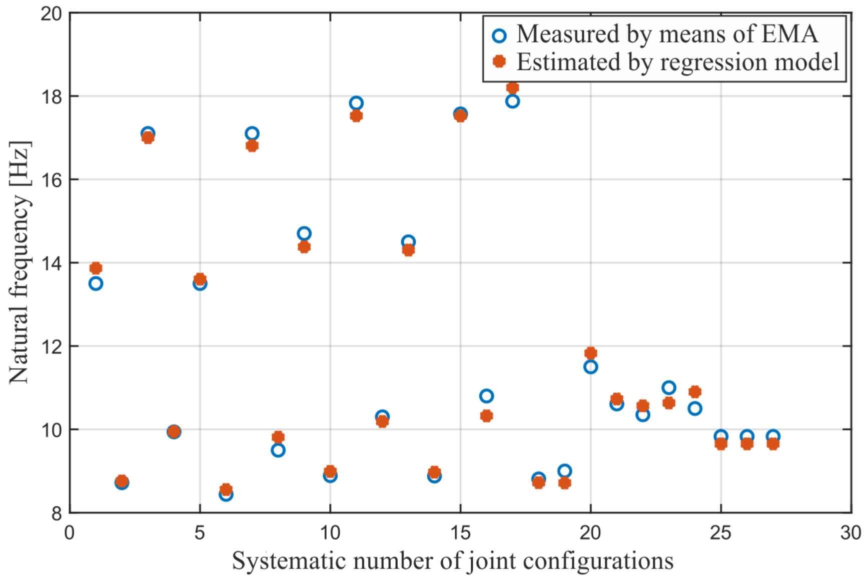


Figure 2: Comparison of the determined natural frequencies between EMA and regression model in different joint configurations

To validate the predictive model, it was compared with a total of 15 points from the real measurement. Figure 2 compares the natural frequencies calculated from the regression model with the results from the EMA. A mean relative error of 3.62% can be observed, allowing a sufficiently good prediction. Thus, a prediction of the natural frequency related to the poses of the robot for the defined milling workspace can be predicted with sufficient accuracy.

4. Results

In order to investigate the effects of the frequency shift, caused by changing the positioning of the milled part and the orientation of the milling spindle, on the surface quality and displacement of the tool, milling tests were run as part of this paper. In addition, the dependence of the milling quality on the milling direction was investigated. Milling on a workspace plane with medium height is expected to be least susceptible to vibrations [16], therefore the investigation of the natural frequencies was limited to a workspace plane at the height of 730 mm. Here, local minima and maxima of the natural frequency were searched for, at milling positions that can be reached by the robot without any problems and that do not lead to collisions with the environment.

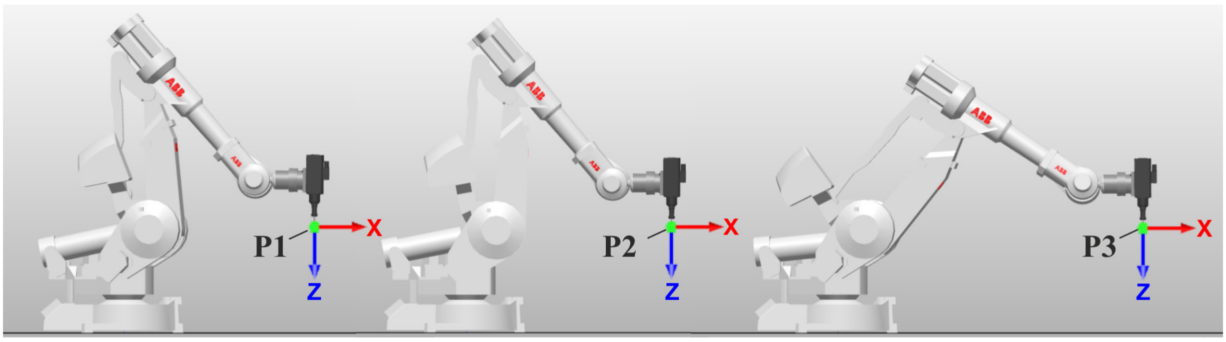
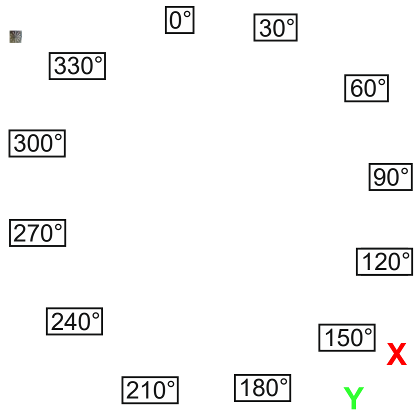


Figure 3: Milling positions in retracted to extended pose (P1: $f = 14{,}4\ Hz, q_2 = -1{,}6°, q_3 = 49{,}0°, q_4 = -4{,}2°, q_5 = -49{,}7°$; P2: $f = 13{,}7\ Hz, q_2 = 5{,}7°, q_3 = 49{,}6°, q_4 = 2{,}5°, q_5 = -48{,}0°$; P3: $f = 10{,}4\ Hz, q_2 = 42{,}4°, q_3 = 33{,}7°, q_4 = -9{,}4°, q_5 = -33{,}0°$)

Accordingly, three milling positions, with orthogonal orientation of the milling spindle to the workpiece surface, posessing a low, medium and high natural frequency were selected; see Figure 3. All milling tests were carried out with a double-edged ball-end cutter with d = 10 mm diameter, since this allows the highest flexibility in terms of redundant robot positions and tool orientations. Subsequently, milling tests were carried out in position P2 (see Figure 3) with different milling spindle orientations. The milling forces were recorded by the *AIT Omega* force-torque sensor. The milling tests were carried out in an aluminum alloy (AlCuMgPb), with the feed rate set to 5 mm/s and a spindle speed of 6,000 rpm.

4.1 Dependency of the milling pattern on the feed direction

In order to investigate the milling pattern in dependence of the feed direction, a star-shaped milling pattern consisting of a total of twelve pockets with identical dimensions was created. These pockets were milled 60 mm long and 10 mm wide. The milling process started in the center of the workpiece surface and ended at the outer end of each pocket at a depth of 6 mm.



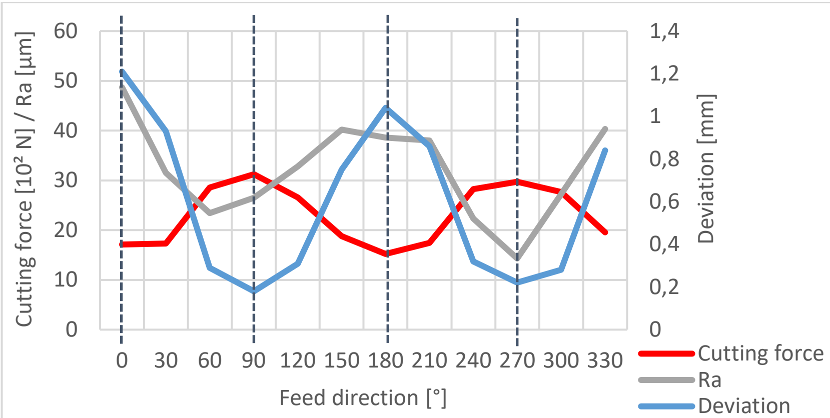


Figure 4: Milling pattern (left) and measured values and milling patterns at position P2 in different feed directions (right)

As expected [16,17], a displacement of the TCP in the rotational direction of the tool was observed; see Figure 4. A significantly lower displacement in feed directions near the y-axes can be seen, whereas the displacement reaches its maximum in the x-axis. This can be explained by the higher section modulus of axes 2, 3 and 5 in the y-direction and the correspondingly acting force direction acting on the robot structure.

As can be seen in Figure 4, this observation is also reflected in the measured milling forces, which are reciprocally proportional to the tool displacement. Due to the higher stiffness of the robot arm when milling in y-direction, higher forces can be applied by the robot, thus reducing tool displacement and improving surface quality. Therefore, the average roughness value R_a in y-direction are about 50% lower than in x-direction.

The milling pattern was carried out at positions P1, P2 and P3, and it was possible to demonstrate correspondingly identical milling behavior in dependence of the feed direction. Thus, a feed in the y direction is clearly preferable to a feed in the x direction. In practice, milling in the y direction cannot always be realized because collisions can occur. Here, a y near milling direction would be conceivable. Furthermore, by using additional axes, an optimization of the milling direction can be realized by the resulting redundant axis configurations in the form of repositioning or reorientation of the robot or the workpiece.

4.2 Dependency of the milling pattern on the position of the milling head

In order to determine the position dependency of the milling results, further milling tests were carried out on the basic poses shown in Figure 3, using the pattern mentioned above. For this purpose, the workpiece was placed at the different positions and the milling tests were carried out with an orientation of the milling spindle orthogonal to the workpiece surface.

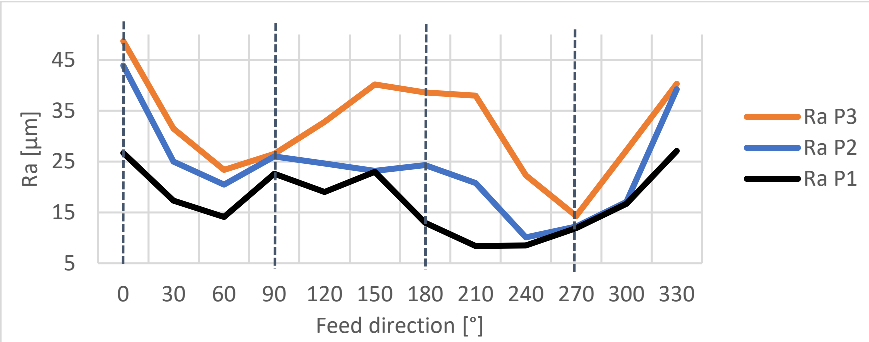


Figure 5: Average roughness value in dependence of the feed direction in the three characteristic positions P1, P2 and P3

Analogous to the milling test in position P2, a parallel sequence of the milling quality in dependence of the feed direction is visible in the two other positions. Nevertheless, a difference in quality can be observed between the individual milling positions. The mean level of the average roughness value $\boldsymbol{R_a}$ changes depending on the position of the robot and the resulting natural frequency. Thus, it can be seen in Figure 5 that $\boldsymbol{R_a}$ averaged over all feed directions is smallest at P1 with 19.5 µm, in the middle range at P2 with 21.3 µm and largest at P3 with 24.2 µm. The natural frequencies therefor are behaving reciprocally proportional to $\boldsymbol{R_a}$. As expected, the displacement of the TCP and the occurring forces also shift with the poses, depending on the natural frequencies. When milling in y-direction, the values for the milling forces are very high and the $\boldsymbol{R_a}$-values are quite low, whereas these values are reciprocally proportional during milling in x-direction. Thus, a correlation between higher natural frequencies, which are depending on the axis configurations, and a worse milling pattern can be seen, while milling with a constant orientation of the milling spindle.

4.3 Dependence of the milling pattern on the orientation of the milling head

As mentioned earlier, a ball-end cutter was used in the experiments, which allows a reorientation in the workpiece surface, while maintaining the same profile shape. This allows a great freedom to make use of the redundancy of the robot, since the orientation of the TCP can be varied. The influence of the different orientations on the milling pattern was tested based on the model in position P2.

The test setup allows a reorientation of up to ±30° in all directions, starting from the initial position P2 shown in Figure 3. The analysis of the milling process in the regression model carried out at the beginning, showed that a reorientation around the RX axis has hardly any influence on the natural frequency. This can be explained by the fact that it is mainly axes 1, 4 and 6 that change here, which have a particularly small influence on the extension of the robot arm. Reorientation around the RZ axis, on the other hand, resulted in a significantly higher change in the natural frequency, since this primarily causes changes in axes 2 and 3. Reorientation around the RY axis had the greatest influence on the natural frequency, since in this case axes 2, 3 and 5 are reoriented primarily, which have a very large influence on the extension of the robot arm. These influences of the change in orientation on the natural frequency were visualized in Figure 6 as an example for one position, based on the regression model. Figure 6 shows possible orientations of the milling spindle with the resulting natural frequency of the robot for this pose. Each orientation of a vector equals the orientation of the milling spindle and the color of the vector equals the natural frequency of the robot. The reorientation results in many redundant poses for the respective orientations, therefore only the respective maximum and minimum natural frequencies were visualized for illustration purposes.

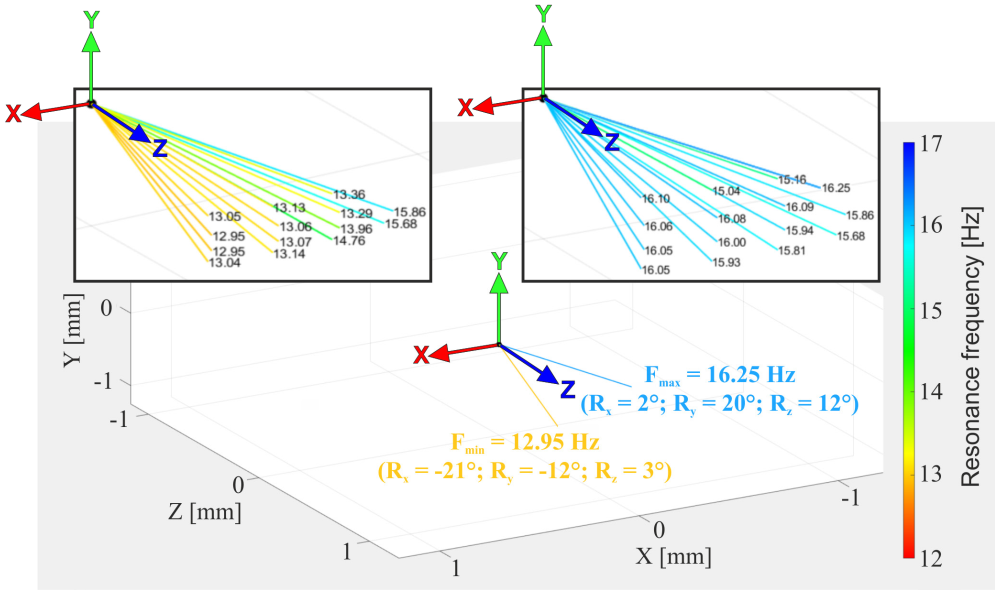


Figure 6: Representation of the reorientation of the milling spindle by ±30° (starting from a TCP zero position of RX = 0, RY = 0, RZ = 0; see position P2) at the pose with the absolute minimum and maximum natural frequencies (center) and respective maximum (top right) and minimum (top left) natural frequencies at the corresponding orientation positions

In order to make this effect of the changed natural frequency as clearly visible as possible in the milling image, milling tests were carried out at position P2 with a reorientation around the RY axis. This reorientation included angles of -30° to +30° starting from the initial position, as shown in Figure 7.

The milling tests were carried out several times and revealed an explicit change in the R_a-values in dependence of the orientation. With the orientation of -30°, where the robot takes up a small extension of the robot arm, a significantly higher average roughness value, than with the orientation +30° can be seen. The best R_a-values could be detected at an orientation of +20° and the worst at -20°.

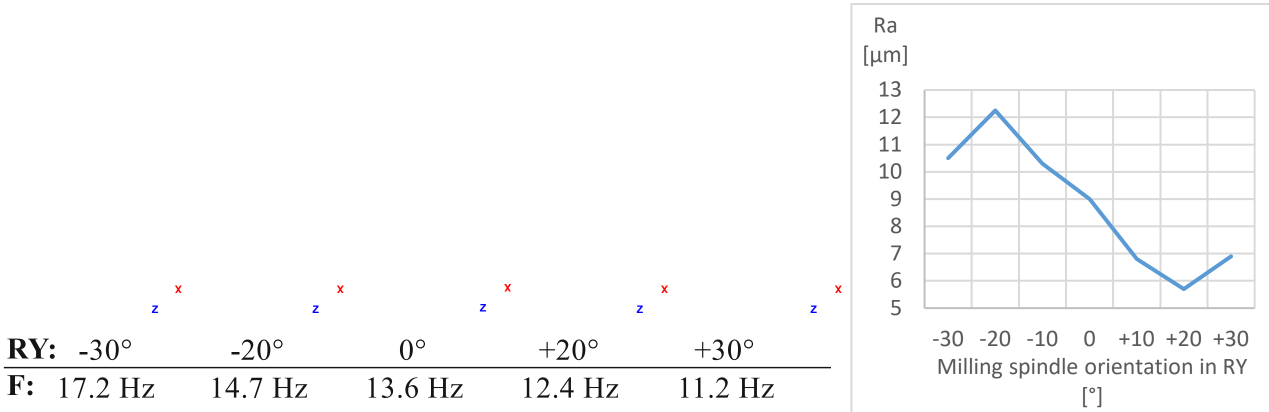


Figure 7: Milling results (left) and R_a-values (right) when reorienting the milling spindle around RY in P2

Contrary to the expectations based on the regression model, a reciprocal proportional relationship between the natural frequencies and the R_a-values cannot be directly proven here. This could be due to various factors influencing the vibration behavior, which have not been considered so far. Among other things, the direction of the force acting during the milling process could play a role, since milling in the direction of or against the robot base results in an opposing load on the motors and joints. Another major influence could be the superposition of the excitation frequencies with the natural frequencies; this effect is to be investigated in more detail in a further paper. It should also be mentioned that due to the design of the ABB IRB 4400, so due to the multi-point suspension of the amplification of the second link, there is an orientation-dependent change in the section modulus and a change in the position of the balance weight. In order to be able to make a statement about the influence of these factors, further investigations will be necessary in the future. However, the milling tests carried out showed a strong dependence of the milling pattern on the orientation of the milling spindle.

5. Conclusion

The high process forces and high-frequency excitation during milling with robots have the effect of displacing the TCP and forming chatter marks, which is still a challenge. Subject to this problem is the low stiffness of the robot and the resulting oscillation in its natural frequencies. The natural frequencies change with the pose of the robot and in this paper could be determined via experimental modal analysis and transmitted to a regression model in order to be assignable to the entire workspace. Experimentally, during the investigations on an ABB IRB 4400, the basic natural frequency was measured in the range of 8.4 Hz to 17.9 Hz. A correlation of the natural frequencies and the milling position was demonstrated in milling tests and a clear advantage of the feed direction in the y-axis was identified. In addition, the effects of the orientation of the milling spindle on the natural frequencies of the robot and on the milling pattern were investigated. Here, a strong dependence of the surface finish on the orientation of the milling spindle could be demonstrated, but not in a causal relationship to the natural frequencies. In order to better understand the existing vibration phenomena depending on the orientation of the milling spindle, further investigations must be carried out and parameters such as the excitation frequencies, the effect of the additional amplification in link 2, and the direction of force must be included in the evaluation.

References

[1] Plyusnina, M., Krewet, C., Rieger, M., Bickendorf, J., Kuhlenkötter, B. (Eds.), 2015. In-Prozess-Erfassung und Auswertung prozessdefinierender Kennwerte beim robotergestützten Polieren von Designoberflächen: Dortmund (12.03.-13.03.2015): Tagungsband. Technische Universität, Dresden, 263-270.

[2] Möllensiep, D., Spörle, F., Störkle, D.D., Thyssen, L., Kuhlenkötter, B., 2019. Robot-based incremental sheet forming to meet current and future automotive industry requirements: Ideen Form geben, 33. Aachener Stahlkolloquium Umformtechnik, 1. Auflage ed. Verlagshaus Mainz GmbH Aachen, Aachen, 379-388.

[3] Trommer, G., 2009. Standardschnittstelle zur Integration von Maschine und Roboter: Einträgliche Symbiose (6), 58–61.

[4] Abele, E., Bauer, J., Stelzer, M., Stryk, O., 2008. Wechselwirkungen von Fräsprozess und Maschinenstruktur am Beispiel des Industrieroboters: Entwicklung von Modellkomponenten für eine Offlinekompensation der spanenden Bearbeitung mit Industrierobotern. Springer VDI Verlag, 733–737.

[5] Glogowski, P., Kuhlenkötter, B., 2016. Schwingungsanalyse redundanter Roboterportale: Schwingungsminimierung im Fräsbearbeitungsprozess durch optimierte Portalstellungen des Roboters. wt Werkstattstechnik online.

[6] Pan, Z., Zhang, H., Zhu, Z., Wang, J., 2006. Chatter analysis of robotic machining process. Journal of Materials Processing Technology 173 (3), 301–309.

[7] Neto, P., Moreira, A.P. (Eds.), 2013. Robotics in smart manufacturing: International workshop, WRSM 2013; co-located with FAIM 2013, Porto, Portugal, June 26-28, 2013; proceedings. Springer, Berlin, 224 pp.

[8] Paucksch, E., Linß, M., 2008. Zerspantechnik: Prozesse, Werkzeuge, Technologien; Vieweg + Teubner, Wiesbaden, 200 - 212.

[9] Azeem, A., Feng, H.-Y., Wang, L., 2004. Simplified and efficient calibration of a mechanistic cutting force model for ball-end milling. International Journal of Machine Tools and Manufacture 44 (2-3), 291–298.

[10] Liu, X., 2009. Machining dynamics: Fundamentals, applications and practices - Machining Dynamics in Milling Processes. Springer, London, 167-231.

[11] Bertini, L., Neri, P., Santus, C., Guglielmo, A., 2017. Automated Experimental Modal Analysis of Bladed Wheels with an Anthropomorphic Robotic Station. Exp Mech 57 (2), 273–285.

[12] Bauer, J., 2011. Methoden der Offline-Bahnkorrektur für die spanende Bearbeitung mit Industrierobotern. Zugl.: Darmstadt, Techn. Univ., Diss., 2011. epubli-Verl., Berlin, 32 - 35.

[13] Basavarajaiah, D. M., Bhamidipati N. M. 2020. Design of experiments and advanced statistical techniques in clinical research , Stringer, Singapore, 100 - 356.

[14] Glogowski, P., Rieger, M., Kuhlenkötter, B., 2016. Eigenfrequenzbestimmung eines redundanten Roboterportals zur Schwingungsminimierung in Bearbeitungsprozessen, 1 - 12.

[15] Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H., Lee, T.-C., 1988. Introduction to the theory and practice of econometrics, 2. ed. ed. Wiley, New York, 839 - 843.

[16] Karim, A., Schmid, S., Verl, A., 2017. 24th International Conference on Production research (ICPR 2017): Posnan, Poland, July 30-August 3, 2017. DEStech Publications, Lancaster, 555 - 561.

[17] Rösch, O., 2015. Steigerung der Arbeitsgenauigkeit bei der Fräsbearbeitung metallischer Werkstoffe mit Industrierobotern. Zugl.: München, Techn. Univ., Fak. für Maschinenwesen, Diss., 2014. Utz, München, 100 - 121.

Biography

Lars Niklas Penczek (*1993), M.Sc. works as a research assistant at the Chair of Production Systems (LPS) at the Ruhr-University Bochum and conducts research in the field of industrial robotics.

Michael Krampe (*1993), M.Sc. works as a research assistant at the Chair of Production Systems (LPS) at the Ruhr-University Bochum and conducts research in the field of industrial robotics.

Till 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 "Production Systems" at the Ruhr-University of Bochum.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Introducing a Decision Support System for Use-Case Specific Object Detection Methods in Production Systems

Kai Müller[1], Stephan-Andrés Saß[1], Christoph Greb[1]
Institut für Textiltechnik (ITA) of RWTH Aachen University, Otto-Blumenthal-Str.1, 52074 Aachen, Germany

Abstract

Object detection has the potential to facilitate the automation of optical quality inspection, achieving a significant reduction of human error. However, the gap between needed expertise to understand and integrate complex object detection systems into production environments and the availability of computer scientists, is hindering its use in the manufacturing industry. A support system for decision-makers unrelated to the subject is therefore required to promote industry utilisation of object detection and effectively manage this otherwise unused opportunity of profitable knowledge. Within the Cluster of Excellency “Internet of Production (IoP)”, such a support system has been developed. Lowering the implementation hurdle of object detection systems is achieved by translating complex information about existing methods into tangible factors, such as quality and cost. In this work we aim to structure relevant object detection techniques and employ a decision tree to provide a user-support based on a use-case-oriented framework. The use-case’s basic conditions and requirements serve as input for the framework. The decision tree gives the suitable object detection method as output, accordingly. For traditional object detection techniques, the characteristics are translated into basic requirements, which are then used as input, e.g., for the template matching method the comparison of a source image with a reference image is translated to the ability to guarantee images for all possible quality deviations. The deep learning (DL) methods are consolidated in the project management triangle consisting of quality, cost, and time. Firstly, an introduction into object detection is given. Secondly, traditional methods are clustered and deep learning methods classified. A description of the decision tree is then presented, before testing results conclude this paper. The developed support system enables decision-makers to evaluate object detection methods for individual use-cases and consequently achieve increased production planning efficiency. The system’s universal design allows for application across manufacturing industries and use-cases.

Keywords

Object Detection; Computer Vision; Deep Learning; Quality Management; Digitalisation; Automation; Internet of Production

1. Introduction

Several quality management tasks in the manufacturing environment are currently performed manually. As fatigue increases and focus decreases, the occurrence of human error implies that a constantly high working standard cannot be guaranteed [1]. Introducing object detection into these use-cases helps to minimise human errors, as machine performance is constant. Historically however, this approach is only an option for enterprises that employ programmers for this specific task and have access to large amounts of computing power and data needed for computer vision tasks. With the increased amount of available computing power, object detection becomes more realistic for most enterprises. However, the recent development in computing power and amount of readily available data has increased object detection possibilities at a higher pace than computer scientists have been educated [2]. This leads to a gap between potential and fulfilled opportunity, due to the scarcity of expertise that is able to perform potential tasks. In this paper we propose a support system based on a decision tree to translate the complex parameters of object detection methods into more tangible factors to support closing this gap.

2. Related Work

Object detection aims to detect instances of objects from given categories and return their spatial location [3]. Historically, so-called traditional object detection methods have been used. With advancements in data obtainability and availability of computing power, several deep learning (DL) methods have been introduced in recent years [4]. The following chapters will give a brief overview and explanation of these two types of methods.

2.1 Traditional object detection methods

Traditional object detection methods can be divided into three stages: *image segmentation*, *feature extraction* and *classification* [5]. Additionally, we consider the *image pre-processing* for a holistic approach that covers the entire process from the captured image to the detected objects. In traditional object detection each of the aforementioned stages is performed separately. The goal of the stages leading up to the classification is to extract features in the form of a feature vector that can consequently be used in the classification to determine which objects are depicted in each image [6]. This happens by observing sets of pixels through a sliding window and moving that window across the entire image. This approach has two limitations: Firstly, each relevant feature needs to be manually selected to perform the classification task. Secondly, these methods become computationally intensive and slow because each image is analysed incrementally window by window [6,7].

2.2 Deep learning-based object detection methods

DL is a subfield of Machine Learning (ML) and employs an Artificial Neural Network (ANN) to recognise and learn the patterns in a given dataset [8]. Images serve as input and the classification as output. To train an ANN a set of labelled data is needed. In the case of object detection, a labelled data set consists of images in which the relevant objects are marked and given the label of the class that this object belongs to [9]. The ANN learns how to classify objects within a given image by comparing the predicted class with the labelled objects in the image and adjusting the biases of the neurons and the weights of the connections between them [8]. In comparison to traditional object detection methods a disadvantage of DL methods is, that they follow a black box model, making the process from input to output difficult to comprehend. However, research is currently being done to improve the transparency of deep learning models [10,11].

3. Proposed concept

The initial aim of the developed support system presented in this paper was to suggest an algorithm or DL model that is best suited for the use-case at hand. After having reviewed fundamental literature, it has become clear, that little to no research has been done on this concept. An overwhelming number of algorithms and DL models represent a barrier for the development of a decision tree comprising the entire spectrum. Instead, in this paper we suggest the foundation for the development of an exhaustive decision support system, by clustering the algorithms and DL models according to the similar characteristics they exhibit. Traditional object detection algorithms can be grouped together based on the input they receive and the output they generate. These groups can further be clustered according to the techniques the algorithms use to achieve the output. A similar approach for DL models is presented in 3.2

3.1 The traditional object detection process

The traditional object detection process is divided into the here defined stages *image pre-processing*, *image segmentation*, the *feature extraction* and the *classification* [5]. Since different algorithms with different parameters apply to each of these four stages, the algorithms are clustered for each stage individually. This way the relevant cluster can be suggested based on the given use-case.

3.1.1 Image pre-processing

Image pre-processing is a processing step in which the image is transferred into a new image. The new image is similar to the original version, but is adapted to facilitate further processing, e.g. by increasing contrast. [12] Image pre-processing algorithms are further clustered according to their functionality into *photometric* and *geometric* image processing. *Photometric* processing uses the brightness of each pixel or the average brightness of adjoined pixels in the original image to calculate the brightness of the pixel in the new image. [12] *Geometric* processing keeps the grayscale value constant, but changes the coordinates of the pixels in the original image to form the new image [13].

3.1.2 Image segmentation

In the stage of image segmentation, regions of interest are defined and isolated [12]. Regions of interest could be the contour of a borehole or letters of the alphabet. This stage can be clustered into four types of segmentation algorithms: *grayscale segmentation*, *contour tracing*, *template matching* and *corner detection*. [7] The *grayscale segmentation* algorithms rely on thresholds of grayscale values. Through the calculation or pre-set of a threshold value, the processed image is turned into a binary image. This means that objects are isolated according to whether they are above or below a pre-set grayscale value. [7] The contour of an object is the organised order of the objects edge points. *Contour tracing* is thus well suited for the detection of large objects with high contrast. [7] In *Template Matching*, a given template from an image block is searched for in a larger target image with the prerequisite that the template may only experience translation or rotation to fit the object in the target image [7]. Template Matching can therefore be used for quality control, if the analysed product must fit a certain template [14]. An *edge detector* searches for the specific location of an edge when the region of that edge is clear. An edge is located according to its direction, height and length, where the direction is defined as whether the grayscale value increases or decreases in the direction of search, the height is difference in grayscale value between neighbouring pixels and the length is described as the number of pixels between which the set height has to occur. [12]

3.1.3 Feature extraction

After having isolated the regions-of-interest, characteristic features are calculated. This process is called feature extraction. During feature extraction, a set of descriptors are computed that help to classify and label objects into categories. [15] Such a set is known as a feature vector. Based on the types of objects isolated, feature extraction algorithms can be clustered into two types: *area-based* features and *contour-ba*sed features. *Area-based algorithms*: These algorithms extract information related to the area of the object. An example for this is the area itself. In some cases, however, the area might not be as relevant as for example the centre of gravity, a characteristic related to the area of an object. [12] *Contour-based* algorithms: Algorithms belonging to this group analyse the contour of the objects. Similar to the area-based approach, information about characteristics such as the centre of gravity can be gathered. However, only the contour is considered in this calculation. The results are of comparable quality to the area-based approach but achieved with less computing power. [12]

3.1.4 Classification

Classification is the central process step in object detection and can be described as the mapping of features from a continuous feature space onto a discrete class space. Examples for classification tasks are the recognition of graphic characters in a segmented image and the designation of a screw in a sorting system with different types of screws according to length and width of the screw head. According to the specific use-case, classification algorithms can be clustered into two categories: *multi-reference* classifiers and *functional* classifiers. [12] *Multi-reference classifiers* need a certain number of prototype patterns for each class. To classify an object, the calculated feature vector is compared to the prototype patterns and assigned to the class with the highest similarity. [12] *Functional classifiers* similar to multi-reference classifiers use classified training patterns. However, these prototypes are not used as patterns, but rather do the functional classifiers attempt to approximate the decision function that maps the feature vector from the feature space to the class space. [12]

3.2 Classification of deep learning methods

DL object detection methods follow a more holistic approach, in which the object detection is done from inputting an image to classifying and localising an object in one or two steps, depending on the chosen type of algorithm. Thus, a clustering of the methods as seen for the traditional methods is not of use here. Instead, the DL methods are clustered according to their functionality into One-Stage Frameworks, Two-Stage Frameworks and object detection methods that are offered by online services such as Google's AutoML or Microsoft Azure's Custom Vision. In order for a subject unrelated person to be able to choose, which of these three types of methods is most suitable for their use-case, the methods are assessed according to the project management triangle consisting of quality, cost and time.

3.3 The project management triangle

The aim of the developed support system is to translate the complexity of object detection methods into factors that are assessable by decision makers unrelated to the field of computer vision. The project management triangle shows, which factors characterise a project [16]. The factors represented by this triangle are quality, cost, and time [17]. Since these are characteristics by which projects are measured, they must be tangible and thus represent an adequate target system for the evaluation of deep learning methods.

3.4 Development of an evaluation model for deep learning-based object detection methods

The factors *quality*, *cost*, and *time* with regards to deep learning-based object detection methods are defined as follows:

Quality: The quality of a DL-based object detection method is represented by the mean average precision (mAP). The most common metric to assess the performance of DL-based object detection methods is average precision (AP) [3]. However, since the AP is computed for each object category separately, the mAP is adopted as a final measure of performance, as this is calculated across all object categories [3,18]. The quality spectrum is continuous, where a high mAP value suggests a high quality.

Cost: The cost it takes to implement a DL-based object detection method is derived from the labour cost of the developer and the computing cost to train the model. As these are highly variable factors, in this paper the cost is defined as binary, where a method is classified as high cost, when a developer is needed to program a model and train it, and classified as low cost, when a cloud service can be used to train a model externally.

Time: In this paper time is also defined as a binary value, differentiating between a method capable of real-time object detection and non-real-time object detection.

3.5 Method clusters and their features

Deep Learning methods are typically divided into two types of frameworks: *One-Stage-Frameworks* and *Two-Stage-Frameworks* [19]. This paper also suggests and includes a third type of framework, that does not fit into the typical groups: *Cloud Services*. The following chapter presents these three frameworks and how they are mapped onto the project management triangle.

One-Stage Frameworks or Single-Stage-Detectors (SSD) are DL-based object detection methods, that perform the object detection in one single step with one artificial neural network [20]. Historically, these SSDs lagged in performance compared to the upcoming Two-Stage-Frameworks [21]. However, since the localisation and classification are done in a single step, SSDs reach very high processed frame rates and are thus eligible for real-time applications [19]. The methods belonging to this group of frameworks are therefore suggested for use-cases, where real-time image processing is a basic condition. Examples for this framework include: RetinaNet, YOLO (v1-v3), Single-Shot Detection and Gradient Harmonising Mechanism (GHM) [21,22].

Two-Stage-Frameworks consist of two separate stages, where in the first stage, so-called candidate proposals that show all objects within the image are generated and most non-relevant locations are filtered out. In the second stage, these candidates are then classified. [21,19] The run-through of two separate ANN means that the framerate is lower than that of SSDs, making real-time processing non-viable [23]. When comparing the mAP of SSDs with that of Two-Stage-Frameworks, it can be seen, that Two-Stage Frameworks operate at a higher quality than SSDs [3]. Examples for Two-Stage-Frameworks include R-CNN, Fast R-CNN, Faster R-CNN, Grid R-CNN, and Double-Head-R-CNN [21,22].

The final option of DL-based object detection is represented by Cloud Services. Building a high-quality DL-based object detection model requires expertise. Cloud Services like Google AutoML, Google LLM, Mountain View, USA or Microsoft Azure Custom Vision, Microsoft Corporation, Redmond, USA are introduced to the market to support people with little or no ML knowledge [24]. These services do not require a developer with expertise but can be executed by using the online service that is offered as a subscription model. The omittance of a developer and its replacement by a temporal subscription to the service result in this being the most cost-effective option out of the frameworks. According to Google, their service also includes real-time image processing support [25]. However, the nature of an online service bares the drawback, that it cannot be used if sensible data is involved in training the model, as this data may be compromised. Even though these cloud services do not represent a certain ANN architecture, it is still considered here, as they represent an equally viable solution along with One- and Two-Stage-Frameworks.

3.5 Description and visualisation of the developed support system

The first step towards finding the most suitable object detection method is to determine whether a deep learning-based or a traditional approach is applicable for the specific use-case, by employing the decision tree developed at Institut für Textiltechnik (ITA) at RWTH Aachen University [26]. A classification method for the individual use-cases may be taken from this paper as well.
Once it has been determined whether a traditional method or a DL method is more suitable, the specific clusters of algorithms need to be established. The traditional approach follows the process steps of object detection, suggesting which cluster is most suitable depending on what type of object is to be detected. The decision tree, as seen in Figure 1, is set up so that the user can plug in the data from their specific use-case and will be given a suggestion for the most suitable cluster of methods.

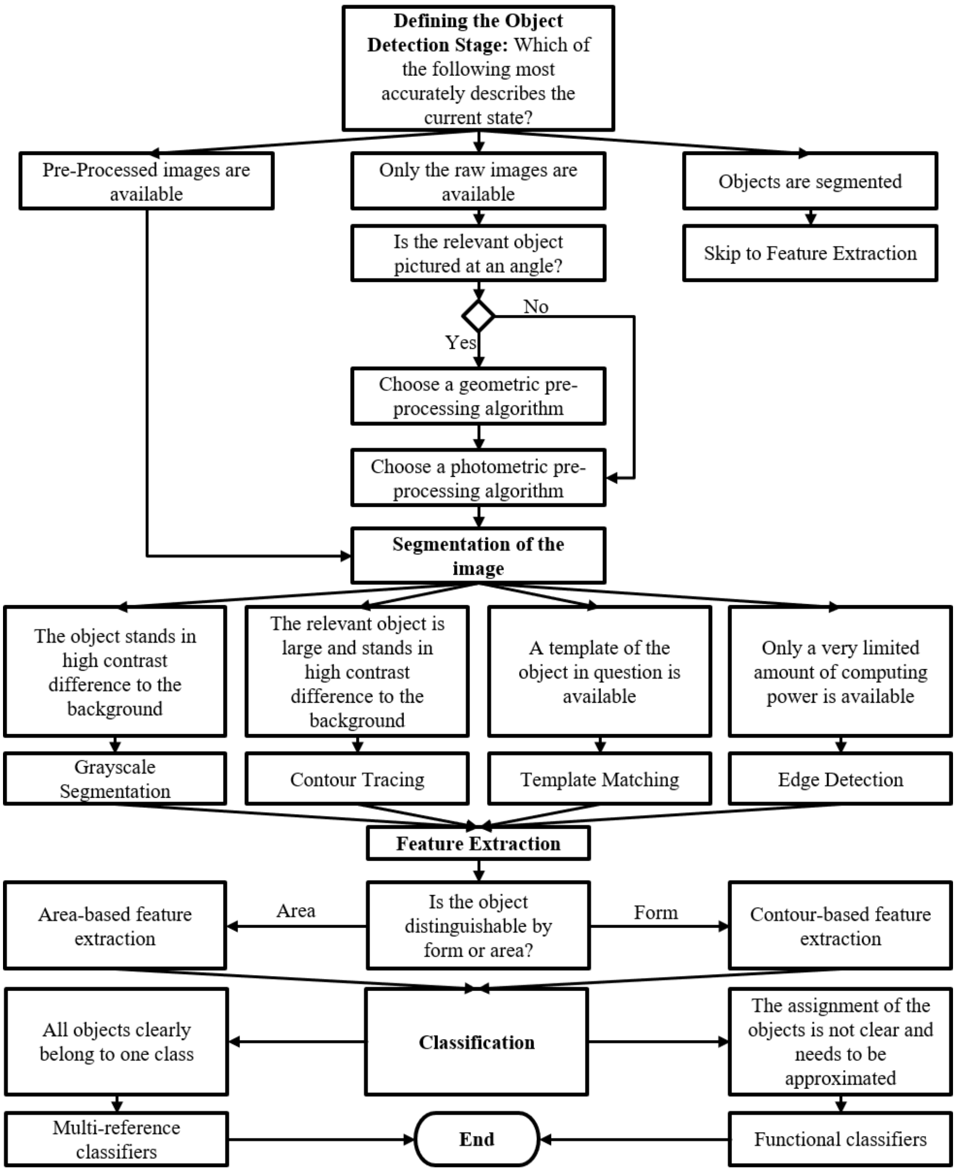


Figure 1: Decision Tree for choosing traditional object detection methods

The decision tree for deep learning-based methods is displayed in Figure 2. The main characteristics by which the cluster of methods is determined are quality, cost, and time. According to these dimensions of the project management triangle, a suitable cluster of methods is suggested for the given use-case. The complexity of the different deep learning-based object detection methods has thus been translated into more tangible factors in which the scope of a project is generally described. It must be noted that there is no concrete outcome for the combination of prioritising a low-cost solution with sensible data. This can be justified with the definition of cost within this paper, as cost is seen as a relative dimension, where the cost of manual labour exceeds that of an online service and thus a cost-effective alternative to the online service would result in a contradiction.

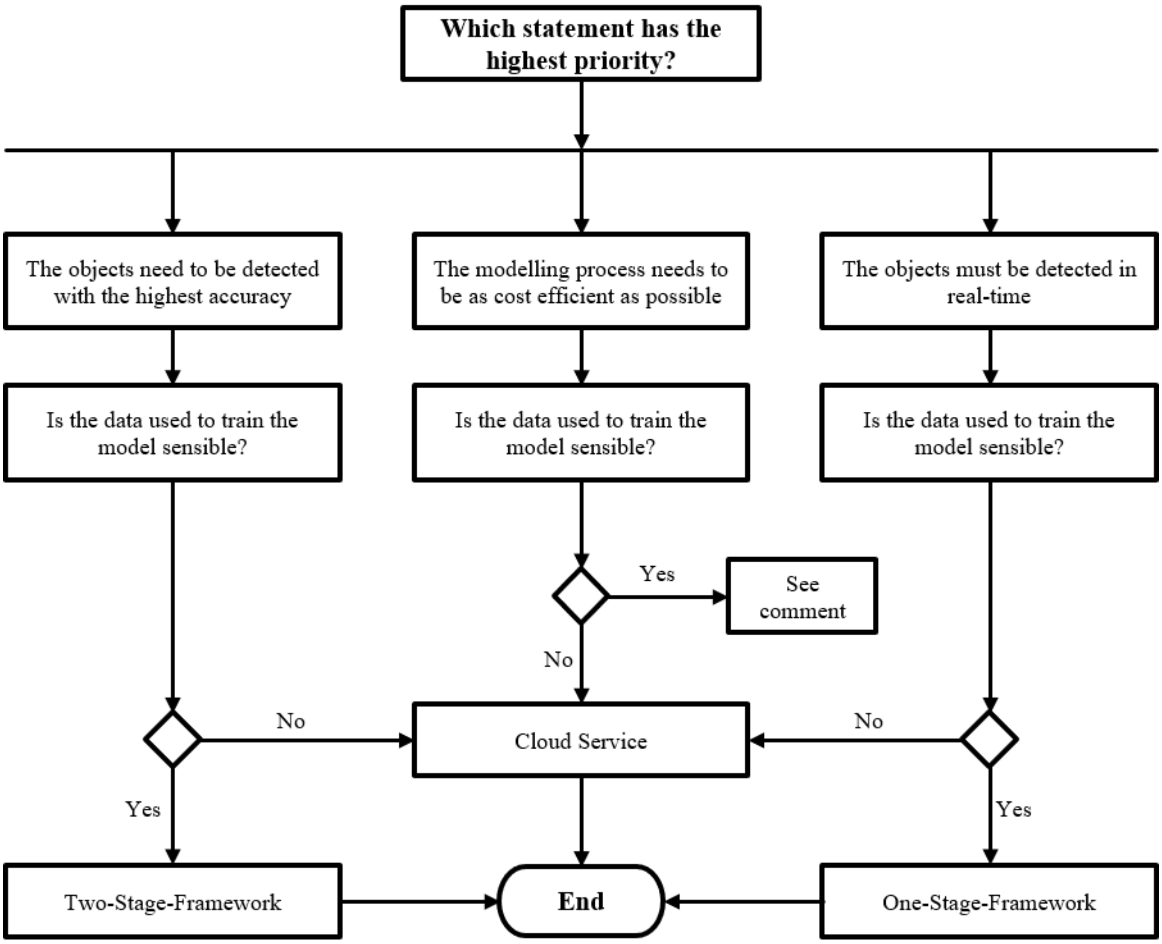


Figure 2: Decision Tree for Deep Learning Methods

4. Use Case

The proposed support system has been applied to a use-case at the Digital Capability Center (DCC) Aachen. In the use-case, the painting of car doors of an OEM was simulated. For this model there is a choice of four lacquers: red, green, blue, and yellow. After the paint has hardened, there are three different outputs: quality okay, white tarnish, and paint runners, resulting in a total of 12 classes that needed to be detected, as seen in Figure 3. The conditions set for this use-case were, that the development of the model had to be cost-efficient, and no real-time object detection was needed.

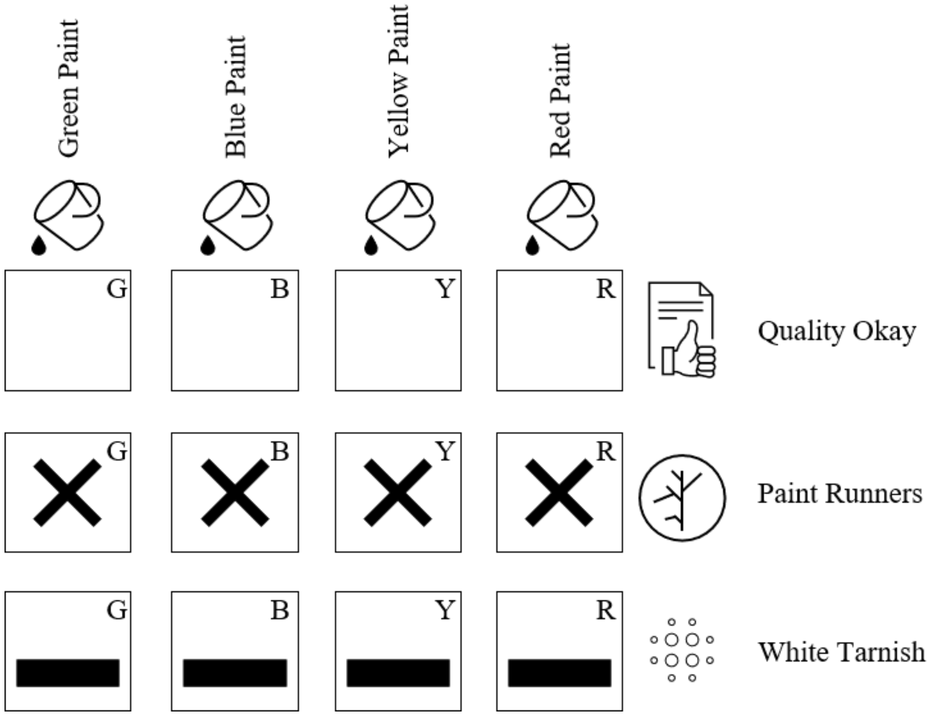


Figure 3: Class types for the DCC use-case

Since this simulation did not encompass client's data, it was not sensible. All objects were readily available at the DCC, thus as much data as needed could be gathered. The availability of high-quality images enabled a deep learning-based approach. Seeing that the data was not sensible, it was possible to use a method offered by an online service. The omittance of a developer and its replacement by a temporal subscription to the service result in this being the most cost-effective option out of the frameworks. With the given parameters of the use-case, the model was then trained with labelled data and Azure's Custom Vision Service. The results of the tested model can be seen in Table 1. To guarantee the comparability of the trained models' performance, the table shows the metrics used by Microsoft Azure, which are commonly used to assess a trained model's performance, where 100% is the optimal outcome for each value. For further information about these metrics, the reader may be referred to [27,28]. This example displays the value of the developed support system. Within a few questions and minimal effort, an object detection method has been suggested, that fits the given use-case and delivered promising results.

Table 1:Testing results for the trained model chosen by the developed decision support system for the DCC use-case

	Precision	Recall	mAP
Value	95.5%	95.5%	98.9%

5. Conclusion and discussion

The application of the decision support system on the DCC use case shows, that with a small number of steps the vast landscape of object detection methods was distilled to a solution that fit the use-case at hand. The support system was able to break down the complex possibilities to solve the problem into a few questions that could be answered by only having knowledge about the use-case and not about the underlying technology. In a next step, to further optimise the presented support system, the clusters need to be granulated further, so that the decision tree outputs not a cluster, but rather a suggestion for an exact algorithm. To optimise the fulfilment of each use-cases potential, this decision tree can be implemented into an existing, user friendly framework developed at ITA [26], that is made to suggest the optimal hardware for a given use-case, thus building an enhanced support system.

Acknowledgements

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder – EXC-2023 Internet of Production – 390621612. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

[1] Myszewski, J.M., 2010. Mathematical model of the occurrence of human error in manufacturing processes. Qual. Reliab. Engng. Int. 26 (8), 845–851.

[2] Jacobs, K., Kuhlmey, A., Greß, S., Klauber, J., Schwinger, A., 2020. Mehr Personal in der Langzeitpflege - aber woher? Springer Open, Berlin, 336 pp.

[3] Liu, L., Ouyang, W., Wang, X., Fieguth, P., Chen, J., Liu, X., Pietikäinen, M., 2020. Deep Learning for Generic Object Detection: A Survey. Int J Comput Vis 128 (2), 261–318.

[4] future. OMahony - Deep Learning vs Traditional Computer Vision.

[5] Zhao, Z.-Q., Zheng, P., Xu, S.-T., Wu, X., 2018. Object Detection with Deep Learning: A Review, 21 pp. http://arxiv.org/pdf/1807.05511v2.

[6] O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D., Walsh, J., 2019. Deep Learning vs Traditional Computer Vision.

[7] Süße, H., Rodner, E., 2014. Bildverarbeitung und Objekterkennung. Springer Fachmedien Wiesbaden, Wiesbaden, 667 pp.

[8] Chollet, F., 2018. Deep learning with Python. Manning Publications Co, Shelter Island New York, xxi, 361 pages.

[9] Suzuki, K., 2011. Artificial Neural Networks: Methodological Advances and Biomedical Applications.

[10] Samek, W., Montavon, G., Lapuschkin, S., Anders, C.J., Muller, K.-R., 2021. Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications. Proc. IEEE 109 (3), 247–278.

[11] Serena Booth, Yilun Zhou, Ankit Shah, Julie Shah. Bayes-TrEx: A Bayesian Sampling Approach to Transparency by Example.

[12] Demant, C., Streicher-Abel, B., Springhoff, A., 2011. Industrielle Bildverarbeitung. Springer Berlin Heidelberg, Berlin, Heidelberg, 420 pp.

[13] Steger, C., Ulrich, M., Wiedemann, C. (Eds.), 2018. Machine vision algorithms and applications, 2nd, completely revised and enlarged edition ed. Wiley-VCH, Weinheim, 494 pp.

[14] Swaroop, P., 2016. An Overview of Various Template Matching Methodologies in Image Processing 2016.

[15] Sklansky, J., 1978. Image Segmentation and Feature Extraction. IEEE Trans. Syst., Man, Cybern. 8 (4), 237–247.

[16] Bohinc, T., 2010. Grundlagen des Projektmanagements: Methoden, Techniken und Tools für Projektleiter, 1. Aufl. ed. GABAL, Offenbach, 186 pp.

[17] Pollack, J., Helm, J., Adler, D., 2018. What is the Iron Triangle, and how has it changed? IJMPB 11 (2), 527–547.

[18] Nguyen, K., Huynh, N.T., Nguyen, P.C., Nguyen, K.-D., Vo, N.D., Nguyen, T.V., 2020. Detecting Objects from Space: An Evaluation of Deep-Learning Modern Approaches. Electronics 9 (4), 583.

[19] Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal Loss for Dense Object Detection, 10 pp. http://arxiv.org/pdf/1708.02002v2.

[20] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2016. SSD: Single Shot MultiBox Detector 9905, 21–37.

[21] Groener, A., Chern, G., Pritt, M., 2019. A Comparison of Deep Learning Object Detection Models for Satellite Imagery, 1–10.

[22] Soviany, P., Ionescu, R.T., 2018. Optimizing the Trade-off between Single-Stage and Two-Stage Object Detectors using Image Difficulty Prediction, 6 pp. http://arxiv.org/pdf/1803.08707v3.

[23] Ale, L., Zhang, N., Li, L., 122018. Road Damage Detection Using RetinaNet, in: 2018 IEEE International Conference on Big Data (Big Data). 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA. 10.12.2018 - 13.12.2018. IEEE, pp. 5197–5200.

[24] He, X., Zhao, K., Chu, X., 2021. AutoML: A Survey of the State-of-the-Art. Knowledge-Based Systems 212 (3), 106622.

[25] Google Cloud, 2021. AutoML Tables Beginner's guide | Google Cloud. https://cloud.google.com/automl-tables/docs/beginners-guide. Accessed 13 March 2021.

[26] Müller, K., Gries, T. Beurteilungskriterien für visuelle Inspektionssysteme, in: , VDI-Z, 162(2020), pp. 59–61.

[27] Henderson, P., Ferrari, V., 2016. End-to-end training of object class detectors for mean average precision, 15 pp. http://arxiv.org/pdf/1607.03476v2.

[28] Padilla, R., Netto, S.L., Da Silva, E.A.B., 72020. A Survey on Performance Metrics for Object-Detection Algorithms, in: 2020 International Conference on Systems, Signals and Image Processing (IWSSIP). 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), Niterói, Brazil. 01/07/2020 - 03/07/2020. IEEE, pp. 237–242.

Biography

Kai Müller M.Sc. (*1992) studied Mechanical Engineering at RWTH Aachen University and started as a researcher at the Institut fuer Textiltechnik (ITA) in 2019. His research is focused on value added management in textile process chains and quality assurance.

Stephan-Andrés Saß (*1992) is a Mechanical Engineering student at RWTH Aachen University and a student researcher at McKinsey & Co.’s Digital Capability Center Aachen. His research is focused on industry solutions in the field of digital manufacturing.

Dr.-Ing. Christoph Greb (*1979) was head of the composites department and member of the Institute's board at the Institute fuer Textiltechnik at RWTH Aachen University from 2013 – 2017. Since 2018 he is Scientific Director at the Institute fuer Textiltechnik at RWTH Aachen University. Furthermore, he is member of the advisory board of the project futureTEX in the programme "Zwanzig20 – Partnerschaft fuer Innovation" of the Federal Ministry of Education and Research (BMBF) and Expert evaluator for the European Commission under the topic CE-FNR-14-2020 "Innovative textiles – reinventing fashion" of the call H2020-FNR-2020. (Horizon 2020 – Food and Natural Resources).

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Development Of A Method To Identify A Suitable Storage, Commissioning And Transport System By Focussing On Automation, Versatility And Costs

Maren Müller, Malte Stonis, Peter Nyhuis
IPH – Institut für Integrierte Produktion Hannover gGmbH, Hannover, Germany

Abstract

Storage planning is an important element of the factory planning and a significant competitive factor in times of an increasing global market [1]. The selection of a suitable storage, commissioning and transport system (sct system) is a major challenge for companies, because of the increasing number of new sct systems with different features. The level of automation and versatility of these systems are intransparent and the required level of both for a certain company is unknown [2,3]. To identify the level of versatility of sct systems a method based on versatility characteristics assigned to the versatility enablers was developed [4]. To determine the required versatility of sct systems for a particular company, a catalogue of change drivers was created. For the level of automation of sct systems, the requirements resulting from product characteristics and performance requirements of the warehouse were identified. The performance of the sct systems depends on the automation level, which can be set by influencing factors such as the degree of digitalization. The required level of automation must be determined by restrictions of the company and the identified possibilities of the systems [1]. At the same time, it is required to consider the costs of the systems as well as their possible combinations. Therefore, to save costs, the aim is also to consider systems which do not fit perfectly to the required versatility and automation level for a company but are still at an acceptable level.

Keywords

Storage, commissioning & transport systems; level of automation and versatility; logistics; selection support

1. Introduction

On the one hand, a huge number of different systems, an increasing number of new features in systems, and the intransparency of the required level of automation and versatility make the selection of a suitable sct system complex. On the other hand, the customer demands for a wide range of variants and a fast delivery is increasing. Hence, an efficient workflow of the logistics processes is essential to exist and survive on the global market [5]. The planning and investment in these systems are rare and of long duration and therefore qualified decisions must be made [6]. The versatility of their systems is a major challenge, especially for small and medium-sized companies (SME's), as this is important to be able to compete [7,8,9]. Existing versatility allows companies to adapt flexibly and quickly to changing conditions, like order variations and individualized products [10]. Furthermore, automation of the systems must be considered [11]. However, the question which individual automation level is useful to handle the storage tasks efficiently is hard to answer, especially for SMEs.

To support companies in their decision-making, a method for selecting and evaluating sct systems has been developed. The procedure and the developed partial methods are signed up in Figure 1.

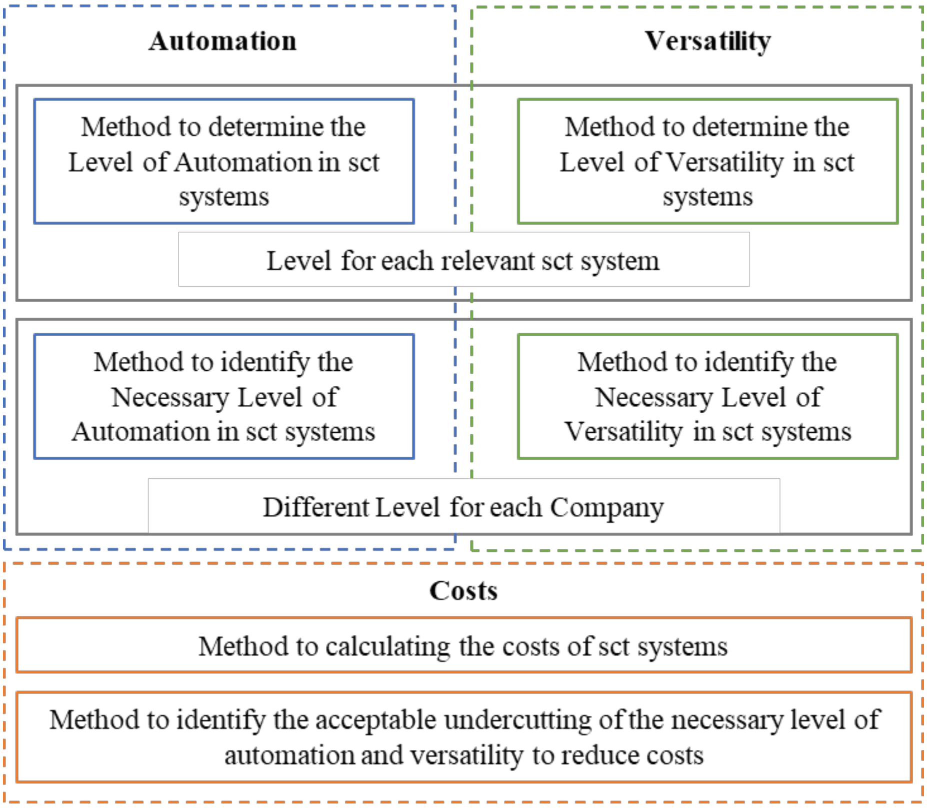


Figure 1: Method to identify a suitable sct system

The following section will first discuss the automation, starting with the method for determining the level of automation in sct systems. Then the method to identify the required level of automation required by the respective company is explained. After that the topic of versatility is elaborated. First the method to determine the level of versality in sct systems is presented, followed by the method to identify the required level for a company. To extend this, costs are considered in the next step. Within the analysis of costs, the first method is about calculating the relevant costs. The aim of the second method is to identify the maximum undercutting of the optimal requirements for automation and versatility, which are accepted to reduce the costs. The integration of costs is done via a filter structure. To support companies in their decision, the results are transferred to a tool, which will be shown next. In the end, the paper gives a conclusion.

2. State of the art

Automation is defined as "the result of automation, i.e. the use of automata" [12]. More precisely, it refers to the "(...) set-up and execution of work and production processes in such a way that humans do not need to be directly active for their sequence, but all processes take place automatically" [13]. The different levels of automation are defined as "the proportion of autonomous functions in the totality of the functions of a system or a technical plant" [12]. Use case specific degrees of automation of different systems can be found in the literature [e.g. 14,15,16,17]. The question of the right degree of automation and the selection of a suitable sct system is unanswered in literature.

A production system is versatile if there are no additional functional units needs if changes influence the storing situation [4]. For the determination of the level of versatility of sct systems and the individual required level of versatility the literature gives two suitable approaches [4,18], which were adapted to the consideration of the sct systems.

Investment appraisal comprises calculation methods used to determine the financial advantageousness of investment projects [19]. A bottom-up procedure, which is described in the literature, was used to estimate the costs of sct systems [20]. The approach to consider the three influencing factors: Automation, Versatility and Cost to select a suitable sct system does not exist in literature yet.

In addition to research in the areas of automation, versatility and cost, research work also exists in *sct systems*, where, for example, the current use of flexible and automated sct systems or their increase is addressed [21,22]. Approaches on performance evaluation (for a selection) also exists [23,24,25,26]. However, these works do not serve to determine the current as well as necessary versatility and automation under consideration of the costs for the selection of a suitable sct system.

3. Design and function of the partial methods in the automation

Automation is the sum of the functions that a system can perform independently, without human intervention. Automation can concern the topics regulation and control [12,13]. In order to identify a suitable sct system, it is required to develop a method to determine the level of automation of sct systems. After this a method to identify the required level of automation individually for each company is described.

3.1 Method to determine the level of automation in sct systems

To determine the level of automation, a difference is made between the categories *mechanization*, *computerization,* and *digitalization*. These categories are elaborated separately to ensure a detailed view [6]. The level of *mechanization* represents the replacement of human muscle power by automation solutions. It was divided into the areas of transport, storage and retrieval and identification of products [2]. Different levels are defined for the individual evaluation of the mechanization. Table 1 shows the classification using the area: transport. The five levels range from “manual work”, which is equivalent to no mechanization, to automatic functions, which correspond to the highest level of mechanization [6].

Table 1: Levels of mechanization “transport”

Category: Mechanization, Area: T*ransport*	Description
0	Manual work
1	Use of additives
2	Machine operation level 1
3	Machine operation level 2
4	Autonomic functions

Computerization indicates the level to which cognitive tasks can be performed with machine support. Computerization is divided into three areas: services before, during and after the transport. “Before transport” describes for example the task of putting the transport units together [29]. The level of computerization in this area range from no support for decision-making to autonomous decisions being made by the system. The area "during transport" is divided into five computerization levels and ranges from manual to autonomous navigation [31]. The area “after transport” deals with the identification of products. The levels range from manual input of information to automatic identification [6].

The level of *digitalization* is divided into four areas: networking at product level, system networking, data processing speed, and data analysis. The area "networking at product level" ranges from manual identification to automatic localization, for example via GPS signals [2,6,31]. In the area "system networking", the lowest level has no communication between the systems. At the highest level, the systems communicate automatically. For “data processing” the level of digitalization goes from no digital support to real-time processing. The fourth area of digitization is “data analysis”, which is divided into four levels based on the data mining context. The levels range from no data analysis to the ability to make own decisions, which is the maximum level of digitalization in this area [2,6,27]. With these levels of automation in the different categories (mechanization, computerization, and digitalization) and their areas the level of automation of each sct system can be determined. If one of these areas does not influence the system types, the area is not included in the calculation. It is also possible that an extra feature could level up a system, especially in the part of digitalization, for example a simple Forklift can be levelled up by an integrated scanner system or a GPS-signal, which has to be notice in the application.

3.2 Method to identify the required level of automation in sct systems

The required level of automation is identified in three steps. In the first step the functional requirements are checked. In the second step the minimum level of automation is determined and in the last step the required performance is defined. Within the framework of the *functional requirements*, the company needs information about his product portfolio: Information about weight and dimensions of the storage unit is required. The *minimum level of automation* is determined according to legal guidelines. Systems that could bring a risk to employees are to be excluded. The company fills in a survey that was developed for risk assessment. For example, information on employees, such as age and gender, is requested. Because this process step only deals with physical support, only the category mechanization levels are considered here. The risk can be assessed by negative points, which were valued for each information [6]. In the third process step the required *performance of the systems* is determined. For this purpose, three performance indicators are identified for each System type. The classification of the performance indicators is shown in table 2.

Table 2: Classification of the performance indicators

Systems	Power indicators
Storage systems	- Storage/removal per hour - Storage utilization factor - Room utilization factor [28,29,30]
Commissioning systems	- Picks per hour - Error rate - Availability [30]
Transport systems	- Transports per hour - Utilization rate - Suitability for the execution of rush orders [30]

Computerization and digitization are also considered in this step. The company selects which level as described in the previous section they want to reach. In this way, a matching with the sct systems can take place, because they have been evaluated in the same way. Systems which do not fit can be excluded [2].

3.3 Matching

These two methods around the topic of automation are merged in the order of the above-mentioned process stages via a filter structure. In the end, the company receives a pool of systems that fulfill the functional requirements, the minimum level of automation and the performance indicators. Under the assumption that a higher level of automation increases the costs, the optimal system in this pool is the one with the lowest level of automation. However, all possible systems are displayed for further evaluations [6].

4. Design and function of the partial methods in the versatility

Versatility describes the potential to carry out organizational, technical, and logistical changes with low investments, with consideration of the interactions of the system elements, and if necessary, in a short time [31]. The procedure for selecting sct systems with regards to versatility is like the procedure for automation. It is required to classify the examined systems and compare them with the required versatility of the company. In the first part the method for determining the versatility of sct systems is presented. In the second part the method to identify the required level of versatility is shown, followed by the resulting matching.

4.1 Method to determine the level of versatility in sct systems

To identify the versatility of sct systems, two valuation approaches are combined and oriented towards sct systems [4,17]. Both procedures result in a catalogue of characteristics defining the versatility enablers. In this case three tables have been developed, under constant consultation with experts, one for each system type. Table 3 shows the characteristics of the versatility enablers for storage systems.

Table 3: Characteristics of versatility enablers for storage systems

Versatility enablers	Characteristics
Universality	- Product and variant flexibility - Nominal capacity etc.
Mobility	- Degree of connection - Spatial mobility
Scalability	- Expandability
Modularity	- System architecture
Compatibility	- Operability - Documentation etc.
Object specific potential for change	- Commissioning

For each characteristic different execution have been identified and evaluated. For example, the characteristic “Product and variant flexibility” has the executions: Not fulfilled, sporadically fulfilled, partially fulfilled, largely fulfilled and fulfilled. In this case the rating is 0, 25, 50, 75 and 100 percent. Every characteristic has its own executions and suitable ratings. Overall, the combination of the different execution ratings for each characteristic resulting in a degree of fulfilment and the percentage of versatility for each enabler and hence also each sct system [6].

4.2 Method to identify the required level of versatility in sct systems

The required level of versatility in sct systems is determined by using a catalogue of change drivers based on the research project “WaProTek” [3]. In this catalogue all change drivers, which can influence the factory objects, are listed. Based on this procedure, the method to identify the required level of versatility for sct systems was developed. First, all relevant driver clusters were identified. In the context of sct systems, the following driver clusters were created: legislators and associations, suppliers, competitors, companies and network, globalization, digitization, employees, and technology. The different change drivers for each driver cluster must be identified. Definitions of the change drivers and related questions about them were formulated. By answering the questions with "does not occur", "occurs occasionally" or "occurs frequently", companies can evaluate the drivers. To assess the required versatility, the next step assigns the various drivers to the versatility enablers, which could counteract the change driver. The drivers can also influence more than one enabler. Table 4 shows one line as an extract of the change driver cataloque.

Table 4: Change Driver Catalogue Extract

Driver cluster	Change drivers	Definition	Question	Answers	Enablers
Suppliers	Changed order quantities	The order quantity describes…	Does your company experience changed order quantities by the supplier?	does not occur	Universality
				occurs occasionally	
				occurs frequently	

For the evaluation of the required versatility by the individual characteristics, the possible answers of the driver questions are weighted in the following dimensions: "does not occur" = 0%, "occurs occasionally" = 50% and "occurs frequently" = 100%. The different drivers and driver clusters are of different importance for determining the required versatility. Therefore, a pairwise comparison is carried out for identifying the weight. This results in the required percentage of versatility in sct systems for an individual company [6].

4.3 Matching

To match the required versatility to the versatility in sct systems, the results can be directly combined. In both methods, the result is a percentage of the versatility enablers, which allows a direct matching [6].

5. Design and function of the partial methods in costs

A survey of companies has shown that in addition to automation and versatility costs are a key factor in the selection of sct systems. This result can be transferred to the entire SME sector [32]. Therefore, costs have to be integrated in the process of selecting sct systems. It has to be analysed whether a cost reduction can be achieved by falling below the identified optimal level of automation and versatility. For this, the first part of this chapter focusses on a method to calculate the costs. In the second part the requirements which allows an undercutting are described.

5.1 Method to calculating the costs of sct systems

The cost estimation was made by a modification of the bottom-up procedure [19,33]. The costs for sct systems were initially determined individually. After that, the total costs were calculated by adding the partial results together. The data input were selected in consultation with the research project-accompanying committee. To make the results comparable all costs must be extrapolated to one year. The following cost drivers influencing the sct systems were determined: Acquisition costs, installation costs, other IT costs, average maintenance and energy costs, number of employees in normal business operations and the depreciation period and method. Acquisition costs, installation costs and other IT costs are non-recurring expenses that arise before the system is put into operation for the first time (installation costs). This includes the costs of the system itself and all other costs such as the cabling, which is required to put the system into operation. There are different procedures for the calculation of depreciation method, which depends on the individual business objective. Straight-line depreciation is chosen for the specific application. This is easy to implement and allows good comparability over the entire life cycle. It is assumed that the systems have no residual value after complete depreciation, are no longer usable, and no disposal costs are incurred. The energy costs and the number of employees required are based on an average utilization of 80%. These cost drivers must be extrapolated to one year to ensure comparability. The average maintenance costs are also calculated to one year. The following calculation is done for each system. Table 4 defines the parameters.

$$\emptyset K_S = \frac{K_A + K_{IN} + K_{IT}}{t_A} + \emptyset K_W + \emptyset K_E + N \times \emptyset K_N \times t_N$$

Table 5: Declaration of the formula

Symbol	Declaration
$\emptyset K_S$	Average costs of the system per year
K_A	Acquisition costs
K_{In}	Installation costs
K_{IT}	Other IT costs
t_A	Amortization period
$\emptyset K_W$	Average maintenance costs per year
$\emptyset K_E$	Average energy costs per year
N	Number of employees in normal business operations
$\emptyset K_N$	Average labour costs per hour
t_N	Working hours per year

For the purpose of this paper it is not required to go into more details of the cost drivers. A view on this level gives enough output about the resulting costs for the use of sct systems. The costs of each system are stored in an individual profile. All cost factors are listed, and the average yearly costs are calculated. This profile will be filled up more and more over the time.

5.2 Method to identify the acceptable undercut of the required level of automation and versatility to reduce costs

First, it must be determined in which areas of automation a lower level is permissible for reducing costs. It was found out that the functional requirements and the requirements for mechanization cannot be undercut. The functional requirements are fixed, because it is not possible to handle a storage unit with not functional fitting systems. The requirements for mechanisation are the second fixed category, because they deal with the health of the employees. A lower automation level is not acceptable in order to not endanger the employees in their daily working tasks. In the cases of computerization, digitization and performance indicators, an undercutting is permissible. Regarding versatility, an undercutting of all versatility enablers is generally allowed. However, the risk of not being able to react extensively enough to future changes in external circumstances must be considered by the companies. Company specific requests (according to the answers in the survey) are used to determine the permissible undercutting. The result is a corridor with systems that are accepted despite the deviation from the optimum. This corridor is based on the company's individual decisions. There are two possibilities for determining the permissible undercut. One option is to define a fixated level of permissible deviation. In the second option each company sets their own limits. In this case, the permissible deviations are made accessible by company-specific queries and thus by individual decisions of the companies. This approach requires an understanding of how the methods are working, the importance of the selection, as well as the possible consequences of the decisions. The advantage is that all company specialties of each company can be accounted for. The query resulted in a corridor on the axes automation and versatility, with systems that are accepted despite deviations from the optimum (Figure 2).

The coordinate system visualises the automation and the versatility executions on the axes. The circles represent the sct systems and the dotted borderlines together with the green area show the optimal level for a particular company. Without the consideration of the costs, the green marked area presents all suitable systems. With the drawn corridor (dotted green rectangle) further systems are considered for the pool of possible systems, with the precondition that these systems would reduce the arising costs compared to all optimal systems. All systems out of the green area and the costs saving corridor are not suitable.

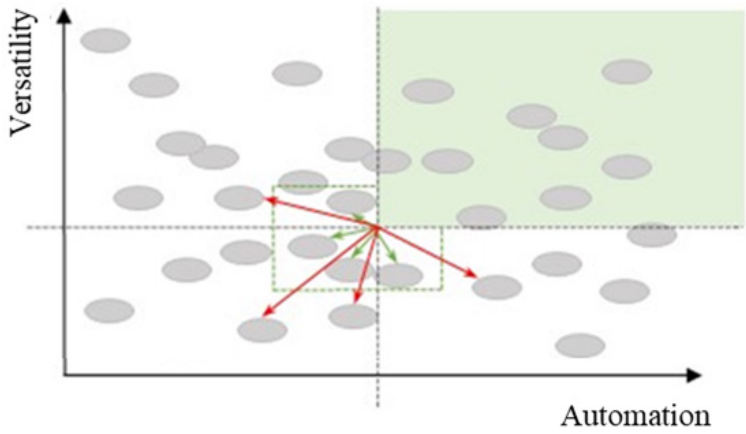


Figure 2: Permitted areas of automation and versatility and corridor for cost savings

6. Method to identify a suitable sct system by focussing on automation, versatility, and costs

For the combination of these partial methods to one overall method, a filter structure is used so that the relevant systems have to go through all the methods of evaluation (Figure 3).

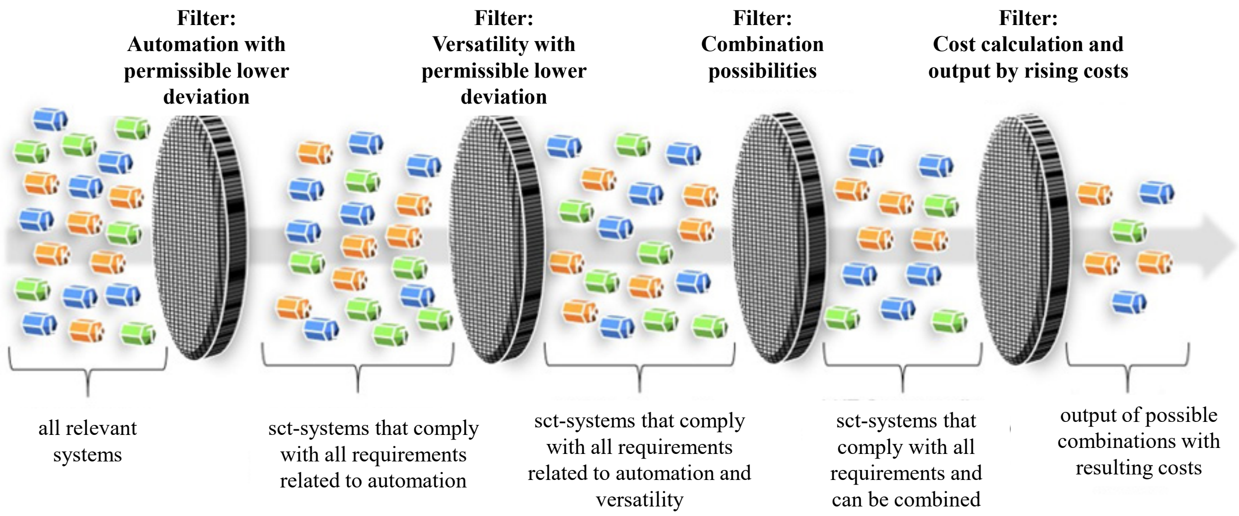


Figure 3: Filter structure

The *first filter* summarizes all the key points of automation. The filter excluded all systems, which are not functional suitable for the storage units. The minimum level of automation includes the mechanization levels and presents the first filter. Systems that do not meet the minimum level of mechanization to be achieved, are excluded. The level of computerization and digitalization are also surveyed when determining the required level of automation. Because some levels can also be achieved through a combination of systems and/or additional products, this has been added. The affected systems, which bring the possibility to reach a level but are not able to do it independently are marked for further evaluation by the company and cannot be excluded here. There is a note that an additional product is required for certain services [6]. Undercutting is possible within the permitted limits in the relevant areas. Information about the undercutting options must be given by the company.

Filter two compares the required characteristics of the versatility enablers with the characteristics of the specific company and the possible undercutting. For the identification of the required versatility, the versatility enablers and the change driver catalogue are used. After answering the questions of the change driver catalogue, all systems are excluded, which do not correspond to the required characteristics of the versatility enablers. Possible undercutting was given again by the company.

In the *third filter*, the systems that are standalone suitable by the selection are examined for combinability. For all possible combinations that are within the optimal range or within the permissible undercutting, the filter is open.

The *last filter* considers the resulting total costs per year. All systems which are in the range of optimal systems depending on automation and versatility automatically pass this filter. The systems, which are in the acceptable corridor to save costs are compared to the cheapest system in the pool of optimal systems. When the costs are lower the systems pass the filter, otherwise they are excluded in this step. The output of the last filter presents the possible sct systems for the specific company.

7. Transfer into an application

The filter structure explained above also corresponds to the query in the application. It starts with the questions about the functional requirements. Here questions about the product portfolio are included. This is followed by the determination of the minimum level and the query of the performance indicators, including questions about the computerization and digitalization requirements and the possible undercutting. After that, the consideration of the versatility is following. The companies must answer questions from the different driver clusters. In the background, the required characteristics of the versatility enablers are calculated in percent and compared with the existing characteristics of the sct systems. The resulting percentage will be shown, and the company can decide if an undercutting is allowed. Then the suitable systems are checked for possible combinations. To answer these questions, it is possible that some companies, especially SMEs will first have to provide the data. Descriptions of the needed data are stored in the tool for this purpose. With this application a high level of prior knowledge and experience is not needed anymore, so SMEs can make qualified investment decisions by themselves. This was programmed using VBA in excel. Via a user interface the company is guided through the questions and receives additional information to answer the questions. When presenting the results, all systems and their characteristics are visible. Thus, the company is given the suggested optimal systems, but can still view all excluded systems and understand why which systems were excluded by the software demonstrator. In addition, the company is provided with descriptions of all sct systems [6].

8. Conclusion

Sct systems and the resulting logistics processes influence the efficiency of companies. In order to support companies in the selection process of a suitable sct system, a method has been developed to select sct systems considering automation, versatility, and costs. In addition to the possibility to determine the required level of versatility and automation and to display the corresponding systems, corridors are determined that show which systems should be considered for cost saving reasons. Thus, possible potential for cost reduction arise, if the companies are ready to do dispense of automation and/or versatility level.

Funding notice

The IGF project 19373N 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 Energy (BMWi) based on a decision of the German Bundestag.

References

[1] Müller, Westbomke, Stonis, 2017. Wandlungsfähigkeit und Automatisierung von Lager-, Kommissionier- und Transportsystemen, ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb, Vol. 112, No. 12.

[2] Müller, Lebbing, Stonis, 2018. Methode zur Bestimmung des Automatisierungsgrads von Lager-, Kommissionier- und Transportsystemen, Logistics Journal, Vol. 2019.

[3] Nyhuis, Deuse, Rehwald, 2013. Wandlungsfähige Produktion – Heute für morgen gestalten, PZH Verlag, pp. 10-21.

[4] Heger, 2007. Bewertung der Wandlungsfähigkeit von Fabrikobjekten. Dissertation Universität Hannover. In: Berichte aus dem IFA, Hannover.

[5] Lanza, Moser, Ruhrmann, 2012. Wandlungstreiber global agierender Produktionsunternehmen. Sammlung, Klassifikation und Quantifizierung. wt Werkstattstechnik online 102 (2012) 4.

[6] Müller, 2020. Schlussbericht zu IGF-Vorhaben Nr. 19373 N, BVL – Bundesvereinigung für Logistik, https://www.bvl.de/files/1951/2125/2131/Abschlussbericht_WALaTra_20200416_IPH.pdf.

[7] Wiegand, Konrad, Jalali Sousanabady, Willats, Deuse, 2014. Quantifizierung von Variabilität durch Big Data Technologien. Gito-Verlag, Dortmund, p. 36.

[8] Guo, Cao, 2013. Strategic flexibility and SME performance in an emerging economy, Journal of Organizational Change Management, 27(2), pp. 273-298.

[9] Alpkan, Yilmaz, Kaya, 2007: Market Orientation and Planning Flexibility in SMEs: Performance Implications and an Empirical Investigation, International Small Business Journal, 25(2), pp. 152-172.

[10] Ludwig, Kotthaus, Stein, Durt, Kurz, Wenz, Doublet, Becker, Pipek, Wulf, 2016. Arbeiten im Mittelstand 4.0 – KMU im Spannungsfeld des digitalen Wandels, Wiesbaden, p. 72.

[11] Nyhuis, Heinen, Reinhart, Rimpau, Abele, Wörn, 2008. Wandlungsfähige Produktionssysteme – Theoretischer Hintergrund zur Wandlungsfähigkeit von Produktionssystemen. wt Werkstattstechnik online, 98 (2008).

[12] DIN IEC 60050-351, 2014. Internationales elektronisches Wörterbuch, Teil 351: Leittechnik 60050-351.

[13] Pollert, Kirchner, Pollert, Bauer, 2016. Duden Wirtschaft von A bis Z. Grundlagenwissen für Schule und Studium, Beruf und Alltag. 6. Auflage. Dudenverlag, Berlin.

[14] Endsley, Kaber, 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, Vol. 42, No. 3, pp. 462-492.

[15] Frohm, Lindström, Bellgran, 2005. A model for parallel levels of automation within manufacturing. In: Pasquino, R (eds.) Proceedings of the 18th International Conference on Production Research, Fisciano, Italy.

[16] Vogel-Heuser, Hompel, Bauernhansl, 2017. Handbuch Industrie 4.0, Bd. 1 Produktion, 2. Auflage, Springer Vieweg, ISBN 978-3-662-45278-3.

[17] Frohm, Lindström, Winroth, Stahre, 2008. Levels of automation in manufacturing, published in Ergonomia – International Journal of Ergonomics and Human Factors, Vol. 30, Issue 3, p. 19.

[18] Heine, 2015. Bewertung des Wandlungspotenzials und Analyse des Wandlungsbedarfs von Kommissioniersystemen. Dissertation Universität Kassel. In: Wenzel S. (Hrsg.): Produktionsorganisation und Fabrikplanung, Band 5, Shaker Verlag, Aachen.

[19] Springer Fachmedien Wiesbaden GmbH, 2019. Gabler Wirtschaftslexikon. 19. Auflage, Wiesbaden, Springer Gabler, ISBN: 9783658195700.

[20] Günthner, Atz, Ulbrich, 2020. Forschungsbericht Integrierte Lagersystemplanung. TU München.

[21] Custodio, Machado, 2019. Flexible automated warehouse: a literature review and an innovative framework, The International Journal of Advandes Manufacturing Technology.

[22] Kamagaew, Stenzel, Nettstrater, ten Hompel, 2011. Concept of Cellular Transport Systems in facility logistics, The 5th International Conference on Automation, Robotics and Applications, https://sci-hub.se/10.1109/ICARA.2011.6144853, Wellington.

[23] Chien-Ming, Gong, de Koster, van Nunen, 2010. A Flexible Evaluative Framework for Order Picking Systems, Production and Operations Management, 19(1), pp. 70-82.

[24] Dallari, Marchet, Melacini, 2008. Design of order picking system, https://link.springer.com/content/pdf/10.1007/s00170-008-1571-9.pdf, Springer Verlag, London.

[25] Staudt, Alpan, Di Mascolo, Rodriguez, 2015. Warehouse performance measurement: a literature review, International Journal of Production Research, 53(18); https://sci-hub.se/https://doi.org/10.1080/00207543.2015.1030466.

[26] Barad, Even Sapir, 2003. Flexibility in logistic systems – modeling and performance evaluation, International Journal of Production Economics, 85(2), pp. 155-170, https://sci-hub.se/https://doi.org/10.1016/S0925-5273(03)00107-5.

[27] Gartner, 2017. Planning Guide for Data and Analytics, https://www.gartner.com/en/documents/3471553/2017-planning-guide-for-data-and-analytics.

[28] Hompel, Heidenblut, 2011. Taschenlexikon Logistik. Abkürzungen, Definitionen und Erläuterungen der wichtigsten Begriffe aus Materialfluss und Logistik, 3. Auflage, Berlin.

[29] Bilcher, Riedel, Schöppach, 2013. Kompakt Edition: Lagerwirtschaft. Grundlagen, Technologien und Verfahren, Wiesbaden.

[30] Pulic, 2015. Lagernutzungsgrad (Flächen-, Raum-, Höhennutzungsgrad), http://www.lagerkennzahlen.de/lagernutzungsgrad.html.

[31] Nyhuis, 2010. Wandlungsfähige Produktionssysteme. Gito-Verlag, Berlin.

[32] ZEW, 2019. ZEW-Pressemitteilung: Kleine und mittlere Unternehmen: Abnehmende Innovationsbudgets behindern Kostenabbau.

[33] Roy, 2003. Cost Engineering: Why, What and How? Cranfield University UK, https://dspace.lib.cranfield.ac.uk/handle/1826/64.

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.

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).

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Low-Level-Code Based Production Model For Improving Material Requirements Planning In ERP Systems

Felix Wimmert[1], Jan Hicking[2]

[1]*Virtual Product Development, University of Duisburg-Essen, Germany;* [2]*Institute for Industrial Management, Aachen, Germany*

Abstract

Single and small-series production companies face specific challenges, such as variable customer order decoupling points (CODP), decreasing quantities and rising cost pressure. This leads to a increasing production complexity and growing requirements on Production Planning and Control (PPC). Digitalization's direct links between objects, people, and machines as well as detailed recording of production progresses opens new solutions for PPC. However, volume of data and the required processing times are increasing. Thus, to achieve near-real-time data processing, a decentralization of decision-making systems can be observed.

The function Material Requirements Planning (MRP) is PPC's original need for Enterprise Resource Planning (ERP) systems. Here, PPC's overall problem (to fulfil primary requirements for products) is divided into subproblems (to fulfil single production orders). Especially companies characterized by an organization in accordance to the workshop principle, high in-house production depth and variable CODP are confronted with high dynamics in their production systems. This ends in significant differences between primary requirements (overall problem) and single production orders (subproblems). Ultimately, these insufficient PPC data result systematically in a non-optimal overall solution despite optimal partial solutions.

This publication combines PPC's fundamentals from existing commonly known models with current implementation concepts of ERP systems. A newly developed Low-Level-Code based Production Model provides explanations for deviations between the overall problem and its subproblems. Furthermore, information flows of PPC can be structured between a periodically actualized vertical and an event driven horizontal information flow. These recognitions lead to an improvement of PPC by ERP systems.

Keywords

PPC; ERP; MRP; Production Regulation

1. Introduction

In recent years production is confronted with increasing market dynamics and demands for individual products. Both result in requests for advanced variability, shorter product life cycles, shorter delivery times and ultimately smaller lot sizes in production [1,2]. In parallel, cost pressure on production is growing because of a worldwide increase in competition [1]. Overall, this results in raising logistics requirements [3] and ask for greater flexibility in production systems [4,5].

PPC's task is optimal positioning of production in conflicting objectives well-known as "dilemma of operations planning". It consists of the opposing objectives of short throughput times, high schedule reliability, low capital commitment costs (low work in progress) and low process costs (high utilization). Therefore, PPC is significant for the success of manufacturing companies [6,3,5]. Digitalization of production systems results in improved data timeliness and data completeness. It creates enormous potential to deal with increased dynamics and complexity [7,8]. However, this does not resolve the specific conflict between customer orientation and logistics costs [4]. Moreover, increasing data in different formats from multiple systems lead to a complexity problem in PPC systems. Availability of valid information is therefore crucial for a successful PPC [8].

In the past, demands for a greater flexibility lead to a decentralization of production systems [9]. It began with primarily organizational development of production into autonomous, self-optimizing, decentralized units under the keyword "lean production" [4]. Nowadays, a further step towards decentralization follows with digitalization [10,2]. Direct communication and event-driven data processing enable a decentralization of decision-making systems and automated control in real time [11,12].

Contrary to the trend of decentralization, successful PPC can only be achieved by a coordinated interaction of different production units [4]. In decentralized decision-making systems, however, it is difficult to coordinate decentralized subproblems and the central overall objective [12]. It therefore remains questionable to what extent separately solved subproblems follow the overall goal. It ends up in two questions: How much decentralized decision-making scope should be given? And how can we ensure technically that the subproblems pursue its overall objective? Consequently, the (search for a) optimum PPC system that combines advantages of real-time decentralized data processing and consistent pursuit of the overall goal has not yet been finally completed.

The objective of this publication is to develop a model that links existing findings of actual PPC models with a current implementation concept in ERP systems. This new model is suitable for explaining relationships between theoretical findings and technical implementation as well as for identifying systematic causes for deviations of decentralized subproblems from their central overall problem.

2. Terms and Definitions

The Hanoverian Supply Chain Model (HaSupMo) [13] is based on the Aachen PPC Model [14] and Lödding's Manufacturing Control Model [4]. It combines the organizational processes of company's supply chain with logistics objectives of PPC. Furthermore, the Two-Stage Control Loop Model [15] illustrates the relation between primary requirements (overall problem) and single production orders (subproblems). Therefore, we combine the Two-stage Control Loop Model with the named and in German-speaking areas commonly known PPC models in section 2.1. PPC's core function of ERP systems is Material Requirements Planning (MRP) according to the MRPII concept [16]. A program-technical method for this is called low-level-code (LLC) procedure [2], which we want to detail in section 2.2.

2.1 Fundamentals of PPC

HaSupMo [13] distinguishes primary requirements initiation into sales order-neutral (production program planning) and sales order-specific (order management). Derived from them, gross and net dependent requirements are determined cyclically by MRP. Net requirements trigger external and internal planned replenishment elements which lead to dependent requirements based on their Bill of Material (BoM) explosion. In accordance with Aachen PPC Model, a distinction is made between external and in-house PPC [14]. Following Lödding's Manufacturing Control Model in task order generation, planned orders are transformed into final production orders. Here, procurement type, lot size and scheduling are finally determined. Further PPC tasks like order release, capacity control and sequencing as well as production itself depend on these production orders [4]. Moreover, in HaSupMo, the task availability check for verification if required resources of a production order are available, is explicitly listed [8].

The Two-Stage Control Loop Model illustrates main information flows of PPC, which are based on variables of a control loop (see Figure 1). Visualized information flows are:

– **Reference variable w_1:** production program planning as well as order management define primary requirements. These are processed in MRP (outer controller) to determine net requirements as well as to generate planned replenishment elements (e.g. planned orders). W_1 operationalizes company's central logistics objectives into the specific fulfilment of demands for products (overall problem).
– **Reference variable w_2:** after order generation final production orders are created. Their planning data (e.g. procurement type, lot size, scheduling) form the production plan, which is input for the inner controller. Hence, production orders of production planning (outer controller) contain reference information for the subproblems.
– **Actuating variable y:** production control specifies *actual input* (availability check, order release) and *actual sequence* (sequencing) of production orders (subproblem) in production (controlled system). Furthermore, production control influences *actual output* of production by capacity control.
– **Control variable x:** With actual data from confirmations of production, Lödding defines control variables *backlog* and *sequence deviation* as difference between planning data of production orders (subproblems) and actual data of production. Moreover, he specifies control variable *work in progress* as difference of actuating variables actual input and actual output [4,1,15,13].

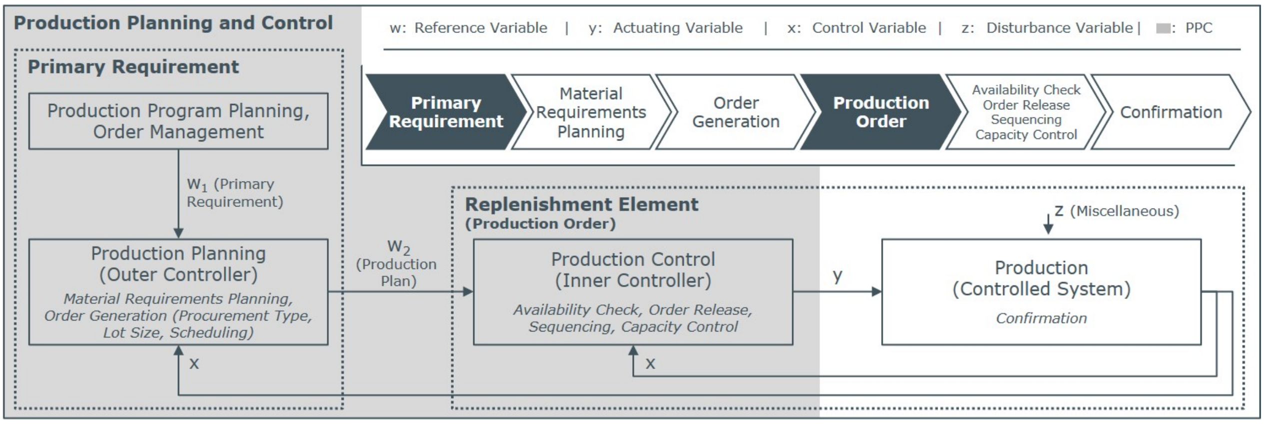


Figure 1: Two-Stage Control Loop Model [1,15]

A classification of different controller types is given by Niehues [1]. Furthermore, HaSupMo and Lödding's Manufacturing Control Model show a direct influence of these control variables on logistics objectives as throughput time, schedule reliability, work in progress and utilization. Due to opposing nature of these objectives, two aspects can be observed: temporal instability and different focuses between functional areas (e.g. purchasing, distribution, production). Hence, an improvement in direction of a company's optimum is difficult [5]. Besides, logistics objectives depend on CODP and are thus subject to market-driven dynamics. For example, in anonymous make-to-stock production, there is a stronger focus on logistics costs (work in progress, utilization). Contrary, in customized make-to-order production, logistics performance (throughput time, schedule reliability) are dominant [5]. However, the transition between make-to-stock and make-to-order production is often dynamic and overlapping. Thus, in production, the conflict occurs not only among different shop floor areas but also between different production orders in one shop floor area [3,5]. Overall, in recent years, there has been a growing focus on logistics performance [2].

2.2 Material Requirements Planning based on LLC Procedure

The Two-Stage Control Loop Model shows transformation of primary requirements for final products into individual replenishment elements (e.g. production orders). It illustrates importance of MRP for PPC (see Figure 1). Central task fulfilling primary requirements of a company [5] are thus divided into subtasks.

Within PPC, ERP systems often serve as a basis. Additionally, they integrate other functions such as accounting, sales and purchasing, and thus form a common data basis. Thanks to central data management, this data can always be retrieved despite different use cases, and redundant data storage is avoidable [17,2]. As in ERP systems many materials are to be processed, MRP takes place successively and hierarchically in stages. A program-technical method for this is called low-level-code (LLC) procedure. Here, materials are grouped according to their lowest BoM explosion level (low-level-code) and processed together according to their ascending LLC (see Figure 2) [18,2]. As a result, all dependent requirements of a material are known at the time it is processed. Thus, it only must be processed once. The required LLC is part of master data of all materials and is usually determined automatically [16]. Figure 2 visualizes the following relations:

- **BoM and LLC (right site, Figure 2):** material's LLC is its lowest BoM explosion level. Purchased part P2 is a component of materials F1 (LLC (1)) and U1 (LLC (2)). Hence, it is on LLC (3).
- **Primary and dependent requirements (left site, Figure 2):** LLC (1) contains only primary requirements, as there are no requirements from above (see LLC (1), Figure 2). On subsequent LLC levels primary and dependent requirements are totaled per material (see LLC (2) and (3), Figure 2).
- **Requirements and replenishment elements (left site, Figure 2):** MRP compares quantities of requirements (negative) and replenishment elements (positive) over time to calculate net requirements. Besides this, both are completely decoupled (see LLC (2), Figure 2).
- **Replenishment element and its dependent requirements (left site, Figure 2):** components of planned orders are determined from material's BoM as well as requirements dates from its routing [18].

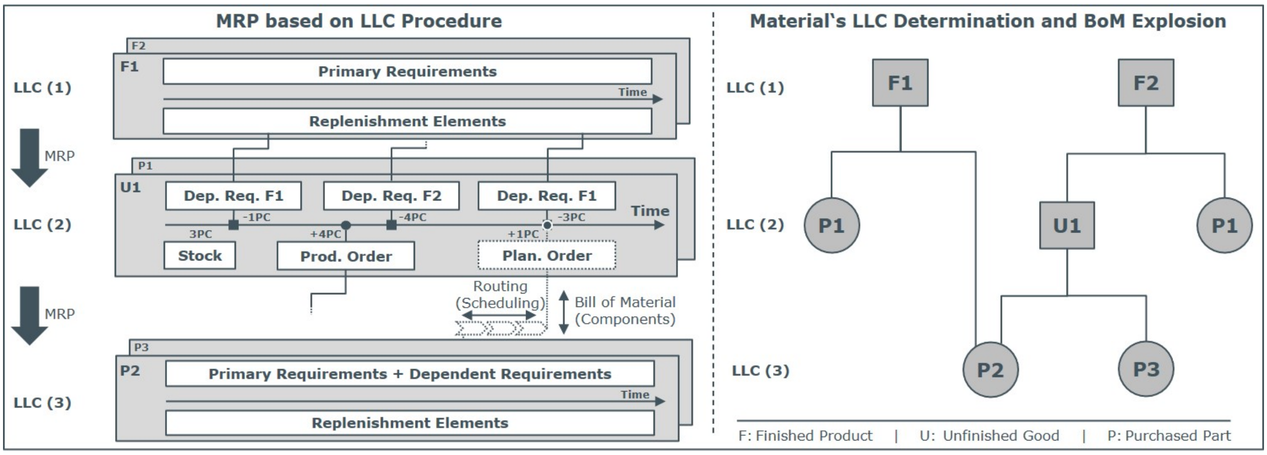


Figure 2: Material Requirements Planning based on LLC Procedure [18]

3. Decision Problem

Every day, production control faces decision problems of which resources are to be used for which production orders. These decisions refer to the production orders (subproblem) and are thus decoupled from their overall problem. In practical use of MRP in several companies with operation type variant manufacturer according to Schuh [14] we determined systematic deviations between planning data of subproblems (w_2) and overall problem (w_1). We observed that especially the following characteristics of companies are critical:

- **Variable CODP:** primary requirements of production program planning are subject to a certain degree of uncertainty due to a lack of a specific order reference. Therefore, primary requirements are not stable. It results in a dynamic in reference variable w_1 and thus significant differences between reference variable w_1 and production orders (w_2).
- **Decentralized workshop production:** undirected material flow of workshop production leads to a further increase of dynamics in production systems and at the same time to a tendency towards high

lead times. Moreover, decentralized organization (responsibility per workshop) favors a pursuit of decentralized goals and focusses even more on reference variable w_2.

- **Complex product structure and high in-house production depth:** a complex product structure coupled with a high in-house production depth leads to complex interrelationships in manufacturing. Therefore, differences between w_1 and w_2 are even more difficult to identify.

PPC's deviations between w_1 and w_2 means that decisions of the inner controller are based on incorrect information. For example, in practical application we observed that in the analyzed use cases on average every production order deviates in scheduling from its requirements at least once. Consequently, this leads to a suboptimal overall result despite optimal solutions of subproblems. Therefore, the question for information technology arises how we can more efficiently support PPC regarding overall logistics objectives. In view of this superior question, the goal of this publication is to develop a model that provides the following answers:

- Explanation for systematic causes of deviations between overall problem and its subproblems.
- Representation of basic information flows from PPC's decisions.

These answers are intended to provide approaches for improvements in practice. They also help us to explain why especially described characteristics of operation type are critical in this context (see section 5). With objective of reducing complexity, we focus on ERP systems with following assumptions:

- Dependent requirements are determined mainly deterministically by MRP in LLC procedure,
- Production planning is organized centrally, while production control is decentralized according to its work center,
- Component availability check is taken as decisive criterion for PPC's task order release.

We justify these as plausible by an operation type (variant manufacturer) and our choice of an established ERP system (SAP ERP).

4. Low-Level-Code based Control Loop Model

We showed that the Two-Stage Control Loop Model (see Figure 1) reflects basic elements of established PPC models. In view of requirements set here, however, it shows considerable uncertainties. On the one hand, it is not obvious what the deviations between w_1 and w_2 are due to. On the other hand, it neither considers details of availability check in production control nor its information flows and technical interrelationships in ERP systems. Hence, we use the Two-Stage Control Loop Model, which is basically suitable for explaining PPC's fundamentals, as a starting point to develop a new model oriented to the questions of this publication.

Challenging in representing relationships between production planning and production control in a common model are their different reference objects. While production planning is based on requirements and replenishment elements (MRP elements), production control as well as production are decentralized, and work center related. We start derivation of a new model with production planning (section 4.1). Afterwards we continue with production control and production (section 4.2).

4.1 Production Planning (MRP element)

LLC procedure processes materials successively according to their LLC. Hence, we conclude that we can better identify weaknesses in a model based on LLC. For this reason, we first detail existing model from Figure 1 in outer control loop according to LLC. Figure 3 visualizes this by separating production planning in vertical dimension. While inner controller, controlled system as well as variables y, x and z remain unchanged compared to Figure 1, reference variable w_1 and w_2 are detailed. We differentiate LLC as follows:

– **LLC (1):** like Figure 2, requirements of materials on LLC (1) include primary requirements only (see input to production planning, Figure 3). Planned and final replenishment elements from in-house production trigger dependent requirements, which are further processed on subsequent LLC. Moreover, production orders form reference variable w_2 for production control (see output from production planning, Figure 3).
– **LLC (n):** here, input to the outer controller consists of the sum of primary and dependent requirements from previous LLC (see input to production planning, Figure 3). As it is the last LLC, no dependent requirements are triggered here (see output from production planning, Figure 3). Usually, only purchased parts are processed on LLC (n).

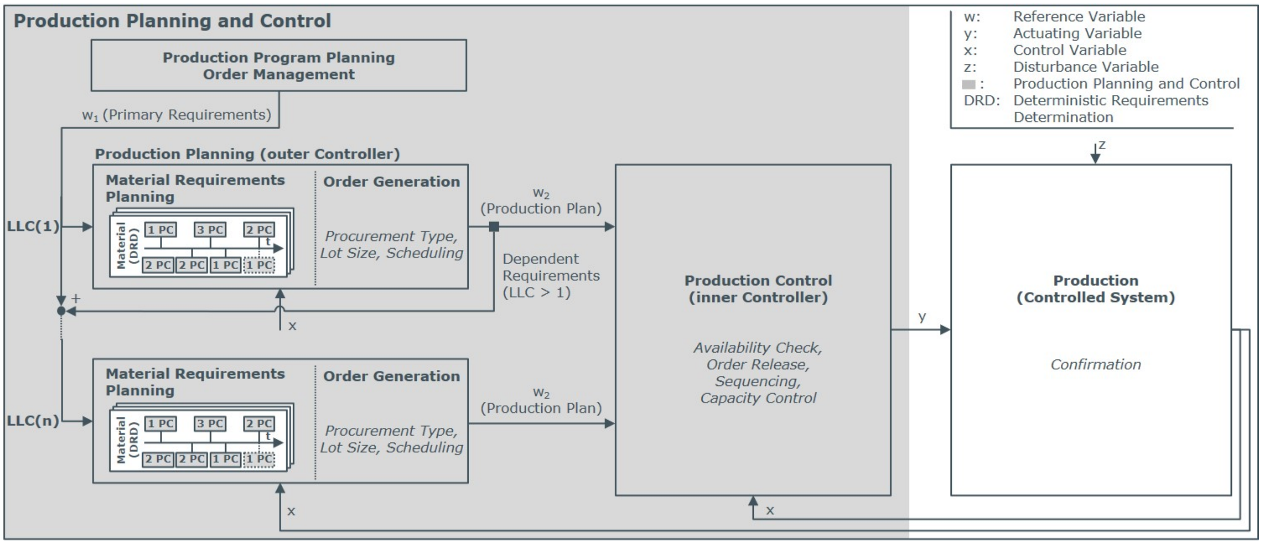


Figure 3 Production Planning in LLC based Control Loop Model

Furthermore, we can derive that a LLC (i) follows the systematic of LLC (1) on its output and LLC (n) on its input side. Finally, Figure 3 visualizes following details:

– Requirements and replenishment elements are decoupled after their processing in MRP. Thus, there is no link between these MRP elements. Hence, production orders are independent from its requirements in production control and production.
– Primary and dependent requirements are summed up per material and processed together on its LLC. Therefore, MRP does not differentiate primary from dependent requirements.

4.2 Production Control (operation of production order)

In this section, we use Figure 3 as a starting point to detail information flows between inner controller and controlled system. As we described previously, the work center related part of a production order is the reference object for this section. In ERP systems, this are production order's operations. They contains all planning data including necessary resources (especially components) for this processing step.

In Figure 3 the vertical dimension is already occupied for differentiation according to LLC. Therefore, Figure 4 divides information flows of production control and production in horizontal dimension according to work centers. Here we retain the original structure of the inner control loop (cf. Figure 1). When taking a more differentiated look at work center B, the following information flows are shown in Figure 4:

– **w_2:** planning data from production planning stored in operations of production orders specify the production plan of individual work centers.
– **Components availability:** corresponding to LLC's definition all components of a production order are on subsequent levels. Thus, to check components availability of a production order requires information from lower LLC.

– **x (1) and y:** according to description of inner control loop in section 2.1.
– **x (2):** a progress confirmation (following open operations) affects production control of next work center, as the production order is ready to be processed there.
– **x (3):** final confirmations (no following open operations) of a production order leads to a change in material stock and thus affects component availability at superior LLC.

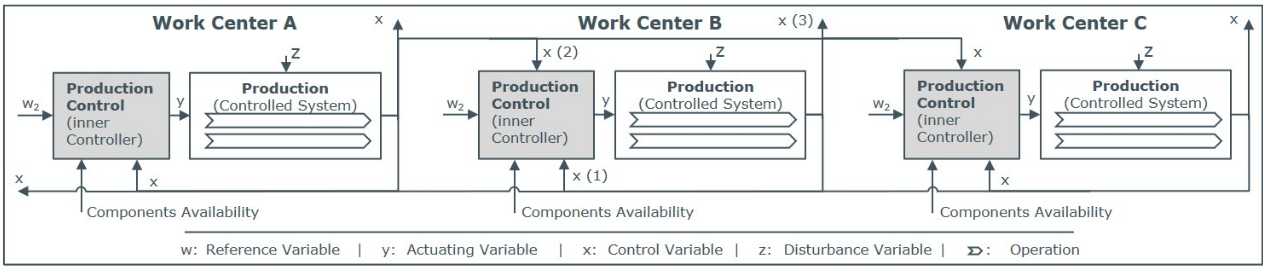


Figure 4: Information flows of inner control loop

Linking these information flows with the model extended in section 4.1, leads to a jointed model shown in Figure 5. There, we took production planning (left side) unchanged from Figure 3, while we derived production control and production (right side) from Figure 4. It requires following further explanations:

– To keep established structure of Figure 1, we had to separate production control in his role as inner controller from its work centers (controlled system). Irrespective of this, the organization of production control remains work center related and thus decentralized.
– LLC structure leads to a differentiation in inner controller and controlled system according to LLC levels. Furthermore, it should be noted that neither production control nor production operates on LLC levels. Dashed lines in Figure 5 indicate that respective worklists are the total across all LLC. Finally, in Figure 6 this becomes completely clear.

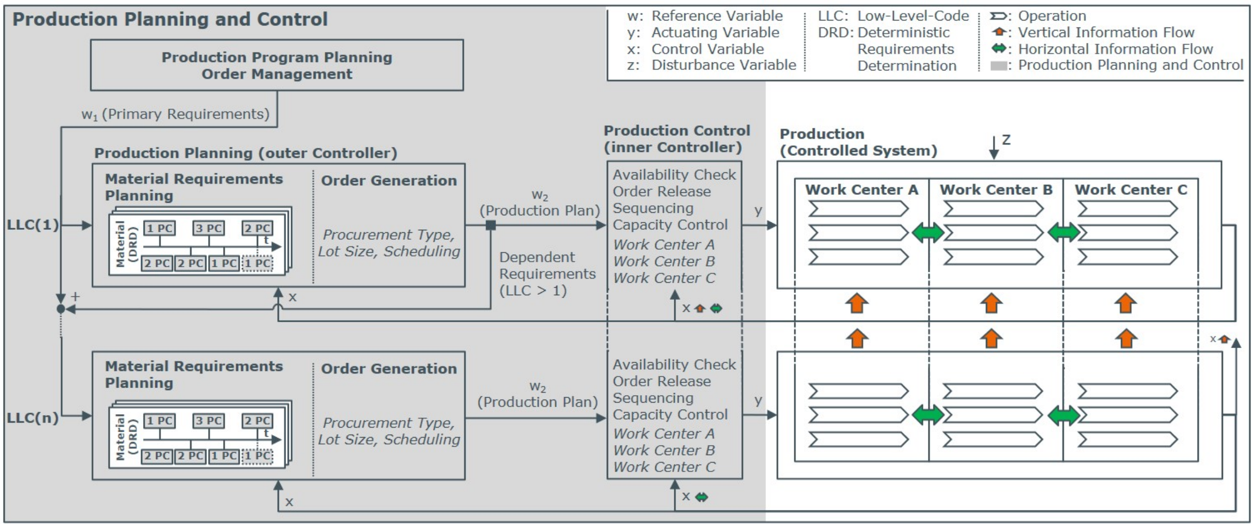


Figure 5: Production Control in LLC based Control Loop Model

Figure 5 shows compared to Figure 3 mainly the same information flows and control variables meanings. Thanks to the differentiation by LLC we can distinguish control variable x of inner control loop by following information flows:

– **Horizontal information flow (horizontal green arrows):** progress confirmations affect production control of next work center (see Figure 5; x (2), Figure 4).
– **Vertical information flow (vertical orange arrows):** final confirmations on low LLC (here LLC (n)) influence PPC's task availability check on higher LLC (here LLC (1)) (see Figure 5; x (3), Figure 4).

In summary, we structured Two-Stage Control Loop Model according to LLC in production planning. It enables us to describe MRP in greater details. Moreover, a closer look on inner control loop makes us to distinguish between a vertical and horizontal information flow. In extension to existing models the new model considers theoretical findings of PPC combined with technical implementation in ERP systems.

5. Application of model

With our new model we can concretely name existing weaknesses in current ERP systems implementation concepts that result in the deviations described in section 3. In Figure 6 we illustrate, in a visualization reduced to one generic LLC (i), weaknesses (1 to 3), which base on the following three findings:

- In MRP constant links between MRP elements are missing. Due to this decoupling, primary requirements are totaled with dependent requirements of higher LLC regardless of their reference to primary requirements. However, with high dynamics in PPC system (variable CODP), planning data of production orders deviate from their primary requirements in terms of quantity as well as scheduling very quickly. Moreover, longer throughput times of workshop production raises duration of decoupling and thus increase these deviations. Hence, these weaknesses in MRP systematically leads on subsequent LLC to growing differences between w_1 and w_2. Furthermore, MRP's single level perspective makes it within high in-house production depth difficult/impossible to identify if there is a deviation in the path above. Once again, this underlines independence of production orders from its primary requirements. Finally, it ends up in the observed deviations between overall problem (w_1) and its subproblems (w_2) as well as faulty information for PPC's decision problem (see 1, Figure 6).
- PPC's vertical information flow is significant for availability check in production control. Due to missing link between requirements and replenishment elements, an availability check is complex. Thus, for decisions which resources are to be used for which orders, both a central-optimal assignment of MRP elements as well as an event-driven data processing for decentralized-flexible decisions are desirable (see 2, Figure 6).
- In addition, information about progress of preceding operations are necessary for optimal production control. This information should be updated event-driven as well (see 3, Figure 6).

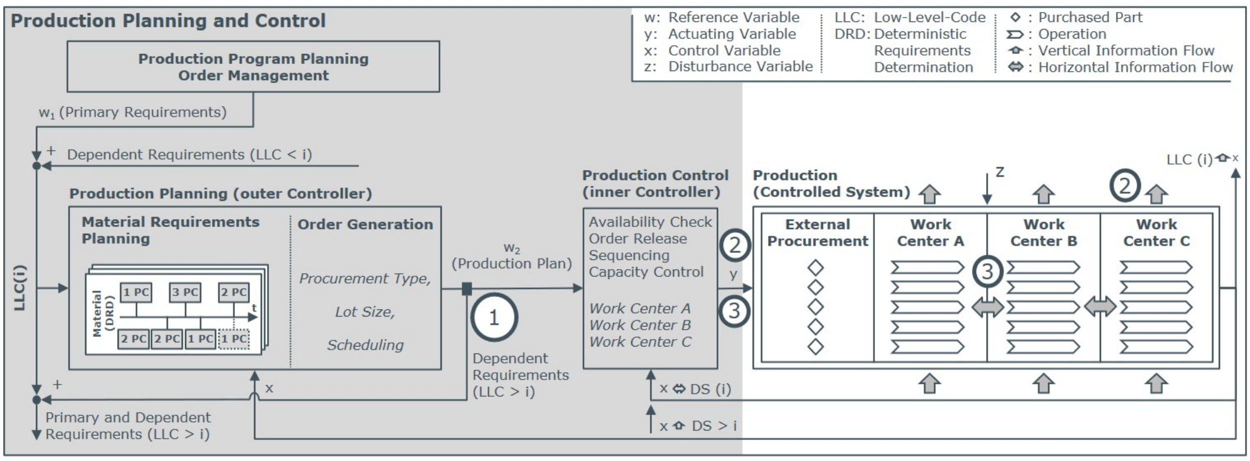


Figure 6: Low-Level-Code based Production Model

In practical application (see section 3), following measures were taken in accordance to named weak points:

- We aligned entire production system with its primary requirements. For this purpose, we adjusted scheduling of replenishment elements to their requirements in MRP. In addition, we saved links of MRP elements temporarily between two MRP calculations. Thus, we could use that information during availability check in production control. It provides a quantity-based allocation of replenishment elements to requirements. Furthermore, we transferred only dependent requirements with reference to primary requirements to subsequent LLC and production control. These measures reduced the deviations between w_1 and w_2 through all levels (see 1, Figure 6).

- In accordance with these changes, only production orders with reference to primary requirements are now available in production control. To maintain decentralized flexibility, we implemented PPC's vertical information flow by a fully integrated warehouse management system. In addition to the quantity-based availability, which is only actualized after a complete MRP calculation, this complemented information for production control by a location-based availability. Moreover, it enabled an event-driven update of location-based availability check. To ensure decentralized flexibility, we used this location-based availability as leading criterion for controlling worklists in production control (see 2, Figure 6).
- At last, we realized PPC's horizontal information flow by integrating confirmation capability of operations (event-driven) as a criterion for its appearance in the worklists (see 3, Figure 6).

In conjunction with these implementations we achieved significant improvements in logistics objectives work in progress and utilization with constant throughput time and schedule reliability. Thus, we reached an improved operating point and research goal of this publication.

6. Summary

In this publication we developed a new LLC based Production Model, which links theoretical findings of PPC with their technical implementation in ERP systems. With this model we can explain root causes of the deviations between the overall problem and its subproblems in MRP. Moreover, it differentiates between a horizontal and vertical information flow.

To resolve the conflict between centralized optimization (overall problem) and decentralized flexibility (subproblems), we propose to correct multilevel scheduling of production orders according to their primary requirement and to assume reference to a primary requirement as a prerequisite for processing production orders in production control. This supports centralized optimization (overall problem). To keep and improve decentralized flexibility (subproblem) we suggest assisting production control in its decisions by an event-driven horizontal and vertical information flow. Furthermore, we recommend this decentralized information flow as leading criteria for the composition of the worklists. For technical implementation we propose to enhance MRP as well as to use a fully in ERP system integrated warehouse management system and confirmation capability of operations. Through these improvements in PPC software, we were able to better synchronize worklists in practical application and better achieve logistics objectives.

References

[1] Niehues, M., Blum, M., Teschemacher, U., Reinhart, G., 2018. Adaptive job shop control based on permanent order sequencing. Prod. Eng. Res. Devel. 12 (1), 65–71.

[2] Wiendahl, H.-P., Wiendahl, H.-H., 2020. Betriebsorganisation für Ingenieure, 9., vollständig überarbeitete Auflage ed. Hanser, München.

[3] Nyhuis, P., Münzberg, B., Kennemann, M., 2009. Configuration and regulation of PPC. Prod. Eng. Res. Devel. 3 (3), 287–294.

[4] Lödding, H., 2013. Handbook of Manufacturing Control. Springer Berlin Heidelberg, Berlin, Heidelberg.

[5] Nyhuis, P., Wiendahl, H.-P., 2009. Fundamentals of Production Logistics. Springer Berlin Heidelberg, Berlin.

[6] Gutenberg, E., 1976. Grundlagen der Betriebswirtschaftslehre. 1. Die Produktion, 22. Aufl. ed. Springer, Berlin.

[7] 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.

[8] Nyhuis, P., Hübner, M., Quirico, M., Schäfers, P., Schmidt, M., 2017. Veränderung in der Produktionsplanung und -steuerung, in: Reinhart, G. (Ed.), Handbuch Industrie 4.0. Carl Hanser Verlag GmbH & Co. KG, München, pp. 31–50.

[9] Thomas, A., Trentesaux, D., Valckenaers, P., 2012. Intelligent distributed production control. J Intell Manuf 23 (6), 2507–2512.

[10] Wiendahl, H.-P., ElMaraghy, 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.

[11] Meissner, H., Aurich, J.C., 2019. Implications of Cyber-Physical Production Systems on Integrated Process Planning and Scheduling. Procedia Manufacturing 28, 167–173.

[12] Scholz-Reiter, B., Freitag, M., 2007. Autonomous Processes in Assembly Systems. CIRP Annals 56 (2), 712–729.

[13] 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.

[14] Schuh, G., Brosze, T., Brandenburg, U., Cuber, S., Schenk, M., Quick, J., Schmidt, C., Helmig, J., Schürmeyer, M., Hering, N., 2012. Grundlagen der Produktionsplanung und -steuerung, in: Schuh, G., Stich, V. (Eds.), Produktionsplanung und -steuerung 2. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 9–293.

[15] Pritschow, G., Wiendahl, H.-P., 1995. Application of Control Theory for Production Logistics - Results of a Joint Project. Ann CIRP 44(1), 421–424.

[16] Stadtler, H., 2015. Purchasing and Material Requirements Planning, in: Stadtler, H., Kilger, C., Meyr, H. (Eds.), Supply Chain Management and Advanced Planning. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 213–224.

[17] Schuh, G., Gottschalk, S., Höhne, T., 2007. High Resolution Production Management. CIRP Annals 56 (1), 439–442.

[18] Vollmann, T.E., Al., E., 2005. Manufacturing Planning and Control Systems for Supply Chain Management, Fifth Edition, Fifth Edition ed. McGraw-Hill, New York, 622 pp.

Biography

Felix Wimmert (*1992) attended a cooperative mechanical engineering education program at University Niederrhein, Krefeld, Germany. After he finished his Master of Science, he started to work for an international company in the ETO to ATO sector as SAP project manager. Nowadays he is responsible for SAP Manufacturing and Logistics at multiple production plants and researching on PPC in context of his PhD studies at University Duisburg-Essen.

Jan Hicking (*1991) studied mechanical engineering and business administration at RWTH Aachen University. After a stint as an energy consultant, he started to work as a digitalization project manager at the FIR at RWTH Aachen University and earned a doctorate in mechanical engineering. Nowadays he is the head of the information management department at FIR. His field of research is the digital transformation in the manufacturing industry.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Comparison of AI-Based Business Models in Manufacturing: Case Studies on Predictive Maintenance

Beatriz Bretones Cassoli[1], Felix Hoffmann[1], Joachim Metternich[1]
[1]Institute for Production Management, Technology and Machine Tools (PTW)/ TU Darmstadt, Darmstadt, Germany

Abstract

Recent advances in Artificial Intelligence extend the boundaries of what machines can do in all industries and business sectors. The economic potential to apply AI in manufacturing results in an increasing number of companies striving to gain a competitive advantage through AI and move into new markets. In this context, particular importance is given to the predictive maintenance of machines. Predictive maintenance promises the possibility of avoiding unexpected machine downtimes and thus increasing the availability of production lines. However, only a few machine manufacturers have a marketable offering of AI-based products or services in their portfolio. Even if technical feasibility is proven, companies lack an understanding of how to integrate AI solutions into new Business Models. This paper thus presents three case studies and their Business Models as examples. Practical considerations and recommendations on the strategical adoption of predictive maintenance technologies are derived.

Keywords

Predictive Maintenance; Artificial Intelligence; Business Models; Machine Tools

1. Introduction

The application of Artificial Intelligence (AI) in production offers great potential for companies along the entire value chain. As a result, global GDP is expected to undergo an increase of up to 14 % or $15.7 trillion by 2030 through the use of this technology [1]. This paper considers the so-called weak (or narrow) AI, which can solve concrete problems [2], for example, by using machine learning algorithms. AI can in this context be described as “intelligent systems created to use data, analysis and observations to perform certain tasks without needing to be programmed to do so” [3]. Predictive Maintenance (PdM) is an example of a task that can be performed using AI technologies. PdM promises the possibility of avoiding unexpected machine downtimes and thus increasing the availability of production plants. The economic potential to apply PdM making use of AI technologies in manufacturing results in an increasing number of companies striving to gain a competitive advantage and/or move into new markets. However, even if technical feasibility is proven, companies lack an orientation to integrate AI solutions into business models [4]. In this paper, the development of AI-driven Business Models (BM) is proposed as the first step to strategical integration. The BM canvas is used as a tool for defining the elements that compose the BM. This work contributes to the existing literature with three example case studies and elucidates common aspects of AI-driven BM for PdM. The focus lies, therefore, on the integration of three AI solutions into the companies' BM.

DOI: https://doi.org/10.15488/11286

This paper is structured as follows, Section 2 provides the state of the art in the fields of PdM and AI-based BM. In Section 3, the case studies are presented. The first case study is from a machine manufacturer that intends to market its PdM solution as an additional service. The second is from a supplier for machine tools who expects to add value to its existing product by implementing a PdM solution through retrofit. Finally, the third case study is a machine tool manufacturer that plans to offer maintenance services inside its load-based leasing model. Section 4 provides a discussion and compares the proposed BMs. Section 5 summarizes the paper and delineates future research directions.

2. State of the Art

This section presents the main concepts and recent developments related to both PdM and AI-based BM. First, the topic of PdM is introduced, highlighting the role of AI, more specifically of machine learning (ML), in the solutions development. Second, BM and BM Innovation (BMI) as well as recent developments of AI-based BM are elucidated.

2.1 Predictive Maintenance

Maintenance is the technical, administrative, and management measures throughout the part's lifecycle to restore or maintain its functional state. Maintenance measures are divided into service and repair. Service measures are for the delay or reduction of the machine's or machine component's wear, while the repair is for restoring the target condition [5]. At present, the maintenance of machines and plants in production is often carried out reactively. Signs of wear are therefore usually not detected until they already restrict the functionality. At this point, however, there is a high probability that delicate machine parts have already been damaged. This form of reactive maintenance often leads to costly repairs and more extended unplanned downtimes. To avoid this, companies carry out maintenance as a preventive measure according to defined intervals (e.g., defined number of cycles, operating hours of the machine). Maintenance in pre-defined time intervals has, however, the disadvantage that, e.g., machine parts that are often still functional and could still be used for a more extended period are replaced prematurely. The use of Predictive Maintenance (PdM) promises to resolve this conflict of objectives [6]. PdM uses the component's condition monitoring data to forecast its state and identify when maintenance is needed, preventing breakdowns and time losses due to unplanned maintenance interventions [7].

PdM approaches are divided into physical model-based, data-driven or hybrid (a mix between these two approaches) [8]. Physical model-based approaches use physical relationships and mathematical representations of the system's failure mechanisms to simulate future states and determine the remaining valid lifetime. These solutions require a profound understanding of the system's characteristics and become unviable for complex systems in which the failure modes and behaviour under adverse operating conditions is unknown. Data-driven approaches use historical and real-time data to forecast the system's state or identify patterns and estimate the remaining valid lifetime [8]. Data-driven approaches use data (time-series, images, and videos) and data analysis techniques, such as statistical methods like exponential smoothing and autoregressive moving average (ARMA) [9], machine learning algorithms like random forest [10], support vector regressor [11] and neural networks [12] [13] to make predictions and estimate when maintenance is necessary.

2.2 AI-based Business Models

A BM is defined as the mechanism that captures value and generates profitable outcomes through the application of a certain technology, it is thus the mediating construct between technology and economic value [14]. A BM is composed of the three dimensions of value proposition, revenue structure and value chain [15] [16]. The value proposition describes what benefits a company promises its customers with a particular product or service. Further, the revenue structure describes how the cost structure and revenue mechanisms in the company are composed so that value is generated from the business. With the value chain dimension, the central processes and competencies required for the implementation of the BM are captured so that the performance or the value proposition can be fulfilled [17]. An AI-based BM is then a business model developed to capture value from AI technologies and applications. In this regard, at least one of the three dimensions of the business model is influenced by the use of AI-methods [18]. On the one hand, existing business models of a company can be transformed through the use of the technology. On the other hand, due to its high disruptive potential, AI offers opportunities to develop completely new business models [19].

Business Model Innovation (BMI) describes the process of developing and continuously improving a company's business model with potential to refine and expand the product and service portfolio [20]. The integration into the corporate strategy and the development of corresponding competencies in dealing with AI topics at all organizational levels is the necessary basis for the development of sustainable benefits from BMI activities in a company [21]. Despite partially divergent objectives and target domains, the corresponding approaches are mostly iterative and range from an initial conception phase to the final deployment of a solution [22] [23]. To carry out each phase of the BMI process, a variety of tools can be used. One of the most widely used tools for the visualization of the BM is the Business Model Canvas (BMC) developed by Osterwalder and Pigneur [24]. The BMC is a framework for visualizing and structuring business models. It is predominantly used at the beginning of the BMI process to analyze the initial situation and to generate initial ideas, as well as providing a holistic overview of the BM components. Based on the three 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 [22]. One drawback of the model for application to AI-based business models is the high degree of generality. Metelskaia et al. address this shortcoming in their extension of the BMC [4]. Based on a comparison of existing approaches to combining BM and AI, they specify the possible content of the individual elements of the canvas. For example, the key partners are extended to include leading IT companies and the revenue streams are extended to include Software-as-a-Service [2].

3. Methodology

The case studies originate from research projects developed with industry partners within the last two years. These research projects were selected because they resulted in proofs-of-concept and initial developments towards an AI-driven PdM solution. The applications, therefore, are not yet being marketed as products. However, the hypothetical BM developments allow comparing the different case studies and serve as guidelines for future implementations. Researchers that worked on the projects completed the BMC from the manufacturer's perspective based on project stakeholders' workshops or project documentation, e.g., project reports. The BMC was chosen because it enables a structured visual representation for the comparison and is familiar with BM developments of other technologies, facilitating the understanding outside of the AI field, as it is industry and BM independent [26]. The fourth section briefly introduces the case studies and presents the BMC.

4. Predictive Maintenance Case Studies

In the following section, three PdM case studies are presented as examples of how to integrate an AI solution into a BM. These case studies illustrate three different types of PdM solutions: the first as an additional maintenance service, the second incorporates not only the developed software but also the necessary retrofitting sensors to enable a PdM solution, the third as an additional service inside a load-based leasing model.

4.1 Machine Manufacturer – Case Study 1

The first case study is from a project developed with an industry partner that manufactures machines for industrial applications. The developed PdM application forecasts a feature representative of a component's health-state and determines when it must be exchanged, allowing planned maintenance interventions at the end of the component's useful lifetime. This application should replace the former reactive maintenance procedure carried out in predefined time intervals.

Figure 1 depicts the technical architecture of the PdM solution. Historical time series data from sensors pre-installed in the machines are used and processed in batches for running this application. The data is transmitted via the OPC UA server to the forecasting application. The forecasting application, hosted in the cloud, receives and pre-processes the data before saving it in a relational database. The regressor retrieves data from the database daily and outputs a 30-day forecast of the target variable. The program verifies whether the forecasted values exceed a pre-defined operation threshold. When the threshold is surpassed, the component must be replaced. The human-machine interface then displays a warning informing the machine operator when maintenance should be performed within the forecasted 30 days.

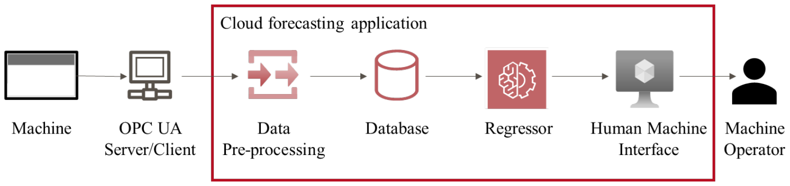


Figure 1. Technical architecture – case study 1

The PdM application is meant to be marketed as an additional service to better plan the necessary maintenance interventions and use the machine components until the end of its useful lifetime. The customer uses the PdM application hosted by the manufacturer, instead of installing the software on its own machines, and accesses the application through the internet. Figure 2 presents the BMC for the first case study.

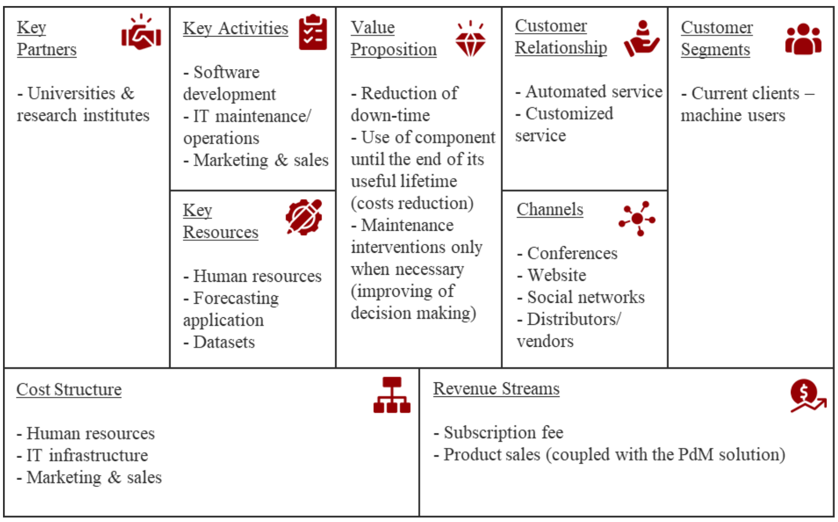


Figure 2. Business model canvas for PdM solution - case study 1

4.2 Supplier for Machine Tools – Case Study 2

The second use case originates from a recently completed research project, which was carried out together with an industrial partner from the supplier sector who manufactures protective covers for machine tools. The project also resulted in a proof of concept as well as an executable functional model on which a later product development can be based. In addition to the development of the technical solution and an AI approach, considerations were also made as to how an initial business model can be designed that is compatible with the prerequisites in the company [27].

As part of the developed PdM approach, sensors record force and acceleration signals at the protective cover during operation of the machine tool. With the help of the recorded data, a feature can be determined that allows conclusions to be drawn about the remaining service life of the component. Based on this, a prediction of the time of failure can be made, enabling maintenance measures to be optimized in terms of time and cost. The solution promises to solve the conflict in industrial practice of a protective cover being replaced too late (reactive maintenance) or too early (preventive maintenance).

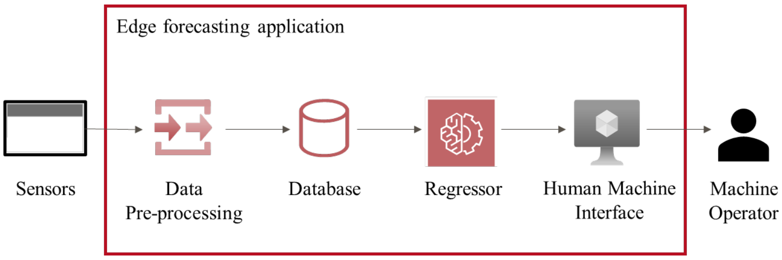


Figure 3. Technical architecture – case study 2

The technical architecture of the PdM solution, as illustrated in Figure 3, is as follows: For the operation of the application, time series data from the installed sensors are collected and pre-processed. This step, as well as making the prediction, is done on an edge device. The regressor continuously receives data and allows conclusions to be drawn about the current wear state at any time. As soon as a defined threshold value for the state value is undercut, the user is notified that the component requires replacement.

In the initial BM, the product is to be marketed as a supplier product directly to machine tool manufacturers via existing sales channels. The company does not plan to provide its own digital PdM service to machine users. This considers two circumstances within the company: Due to the company's origins, there is no expertise in electronics and software development to date. Additionally, it is difficult and time-consuming to build up the necessary expertise because of a persistent shortage of skilled workers. On the other hand, many machine tool manufacturers themselves already have advanced service platforms. The marketing of an own service application without integration into the ecosystem of a manufacturer is not promising. Figure 3 presents the AI BM canvas for the second case study.

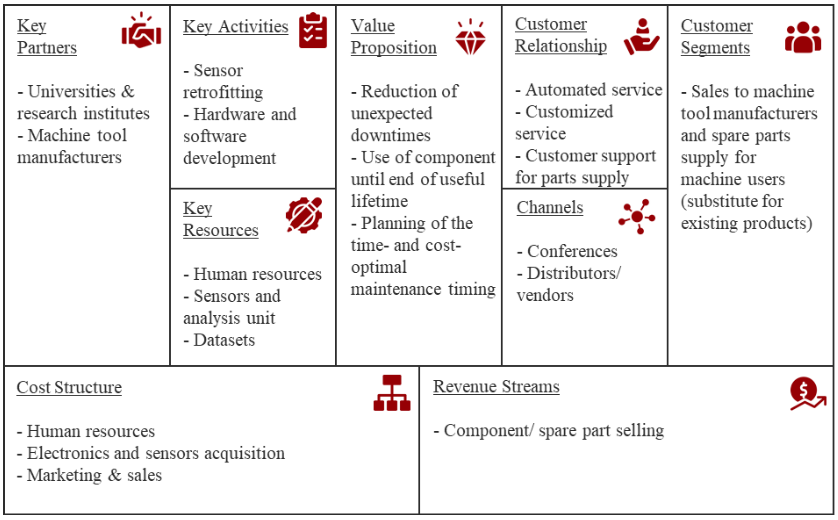


Figure 4. Business model canvas for PdM solution - case study 2

4.3 Machine Tool Manufacturer – Case Study 3

The third case study is derived from an ongoing research project conducted with industry partners. In the project a stress-oriented, data-based payment model for machine tools called “Pay-per-Stress” is developed. Figure 5 provides an overview of the technical architecture. The machine user leases the machine from a lessor who acquires the machine from the machine manufacturer. The payment model has two main components: a monitoring system, which measures the stress linked to the machine and calculates the remaining useful lifetime (RuL), and an incentives system, which sanctions harmful and rewards regular usage, in order to align incentives [28]. The goal of the model is to link stress and actual machine wear and use it as a leasing indicator, mitigating the information asymmetries and inefficiencies from classical leasing and pay-per-x models.

The architecture of the model is visualized in Figure 5 and designed as follows: Sensors installed in the machine are used to record process data during operation and then encoded on the customer’s site. The data is then forwarded via blockchain to the machine manufacturer, who processes the data on its own premises. In the case of maintenance services, the results of the evaluation are forwarded to the customer. This makes it possible to provide the customer with recommendations on how to operate the machines in a load-optimized way and when maintenance measures need to be carried out. The BM in this paper is considered from the perspective of the machine tool manufacturer and is restricted to the PdM service that can be provided based on the RuL estimation.

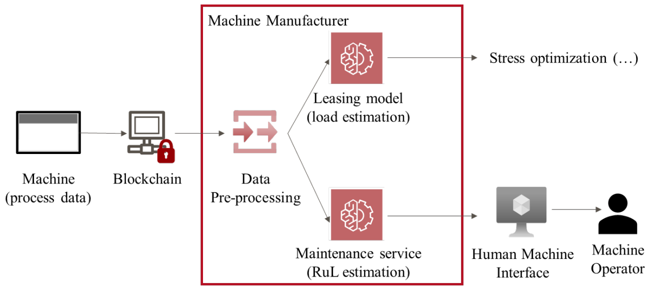


Figure 5. Leasing business model architecture

AI is used as the data-driven technique for calculating the RuL based on process data. On the one hand, the insights gained are used to operate the payment model. On the other hand, it can also be used for the further development of the company's own products and for the operation of maintenance services at the customer's site. Figure 6 presents the BMC for the third case study. The presented BM complements the leasing model BM with additional maintenance services.

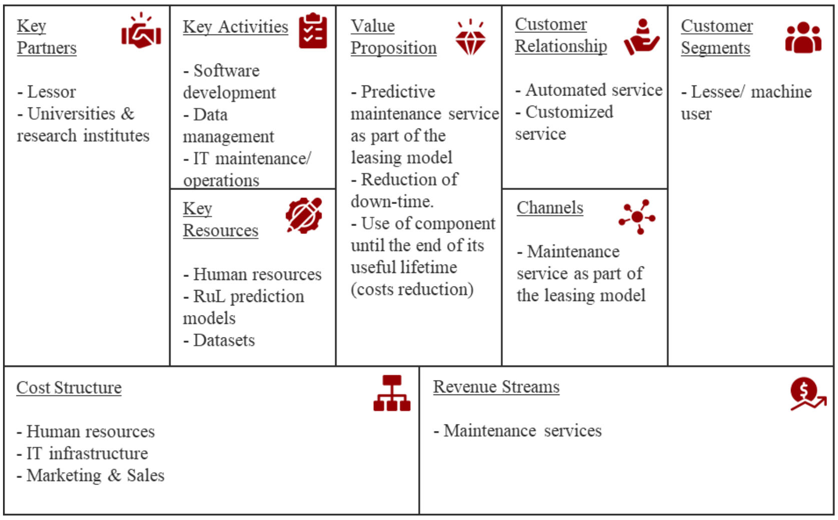


Figure 6. Business model canvas for PdM solution - case study 3

5. Discussion

This section discusses similarities and differences of the presented BMs and highlights opportunities for the strategical adoption of PdM. Table 1 compares the three BMs. Next, insights concerning common aspects to the case studies are presented.

Table 1. Comparison of Business Models

Aspect	Similarities	Differences
Key Partners	- Universities and research institutes.	- Additional key partners for case 2 and 3, since they depend on other partners (machine manufacturers and lessors, respectively) to market the AI solutions.
Key Activities	- Software development. Development of the PdM application.	- Case 2 additionally encompasses the retrofitting of hardware for data collection. Cases 1 and 2 make use of sensors pre-installed in the machines. - Cases 1 and 3 have the IT maintenance and operations. Case 2 uses existing platforms from the machine manufacturer to operate the AI solution.
Key Resources	- Human resources which include the data scientists, software engineers and other members of the AI team [29]. - Datasets/ data. Generated from the machines and production processes and necessary to estimate the components RuL. - AI applications. Including ML algorithms and pipeline.	- Case 2 additionally encompasses the electronics and sensors for retrofitting.
Value Proposition	- PdM objectives: reduction of unexpected downtimes, use of component until the end of its useful lifetime (cost reduction), planning of time- and cost-optimal maintenance measures.	
Customer Relationship	- Automated and customized service. Automated and customized because the customers interact with human-machine interfaces to specify and customize the PdM services. Customer support is necessary to accompany the installation and correct functioning of the application.	- Case 2 foresees customer support for direct sales and part supply. This support is necessary in case of retrofitting of previously acquired machines.
Channels		- In cases 1 and 2 conferences, distributors, and vendors. Channels which are already used in the current BM. - Case 1 additionally foresees advertisement through its website and social network.

		- Case 3 uses as channel the leasing model and offers the PdM solution as an additional service.
Customer Segments	- The new BMs mainly aim at existing customer segments (respectively machine users in case 1 and 3 as well as machine tool manufacturers in use case 2).	- Case 2 foresees an additional opportunity opportunity for direct sales to machine tool users that want to retrofit their machines and allow PdM for the protective covers.
Cost Structure	- Human resources. - Marketing & Sales.	- Cases 1 and 3 also include the IT infrastructure to operate the PdM application. - Case 2 has additional costs for the retrofitting equipment acquisition and does not have additional IT infrastructure costs, since the application would run in the machine manufacturer's platform.
Revenue Streams		- Case 1 through subscription fees and coupled sales with the current products (machines). - Case 2 through sales of the retrofitting equipment for PdM as components/spare parts. - Case 3 through maintenance sales.

It is worth mentioning that the three case studies originate from projects conducted in partnership with research institutes in publicly funded research projects, meaning that it is not a coincidence that for the three use cases, universities and research institutes are identified as key partners in the canvases. This form of collaboration is often used as the first step in creating a technical proof of concept for the content of AI-based BMs. Advantages arise here from the broad expertise of research institutions in innovative technologies, as they deal with these at an early stage. Companies also appreciate the fact that there are often no competing economic interests despite cross-industry collaborations. In addition, cooperation with research institutions is often a necessary prerequisite for gaining access to public funding [30].

Particularly in Germany, where the case studies were carried out, the government offers public funding programs and initiatives aimed at developing AI-based solutions in manufacturing. In addition to the expenses for the personnel required to carry out the project work, expenses for material resources and subcontracts can also be financially supported. Contrary to use cases 1 and 2, not only bilateral consortia are possible, but also consortia consisting of different university and industrial institutions as in use case 3. Public funding allows companies to engage in pre-competitive ventures at a reduced financial risk.

A common aspect to the presented case studies is the central and enabling role of AI technology in the BM (evident in the *Key Resources* field of the BMC). As reinforced by other studies, AI initiatives need to be core to a company's business strategy to create meaningful value and scale [31]. It is argued that the development of a BM for the PdM solutions is the first step for its strategic adoption. Also the *Cost Structure* and *Revenue Streams* need to be aligned with the company's long-term goals in terms of hiring skilled personnel and/or providing training to build an in-house AI team [32]. There is a difference between adopting AI for improving existing business processes, e.g., in manufacturing and in marketing, and adopting AI for developing and enabling new products and new services from which the company can capture value. The second is where the true potential of AI as a competitive advantage lies, and can only be realized when the AI initiatives are developed in-house and are part of the company's strategy [31].

The customer, considered in the fields *Customer Relationship* and *Customer Segments*, is central in the BM development [20]. AI-driven solutions not only have the potential to fulfil customer needs (such as PdM), but also enables new marketing possibilities (described in *Channels*) and image gains, bringing an innovative aspect to the company's portfolio.

6. Summary and Outlook

This paper presents three case studies from the manufacturing industry that have developed AI-based BMs to capture value from PdM solutions. The first case study is from a machine manufacturer that intends to market its solution as an additional service to its machines. The second use case is from a supplier for machine tools who expects to add value to its existing product by implementing a PdM solution through retrofit. Finally, the third case study is a machine tool manufacturer that plans to offer maintenance services inside its load-based leasing model. The business model canvases are presented to identify the main aspects of each case. The discussion section provides a comparison between the case studies, highlighting the similarities and differences and showing that there is not only one option for marketing PdM solutions. AI is perceived here as the enabling technology that allows the development of new BM. Pre-requisites for the AI solutions presented here include not only skilled personnel but also data, in the presented use cases either generated by sensors pre-installed in the machines or through retrofit efforts. This work contributes to the existing literature with three example case studies and elucidates common aspects of AI-driven BM for PdM. Future research includes investigating the actual implementation of the developed BM and identifying the success factors.

Acknowledgements

The research leading to these results has received funding from the German Federal Ministry for Economic Affairs and Energy (Kompetenzzentrum Darmstadt and Pay-per-Stress Project) and from the Digital Ministry of the state of Hessen - Germany (SensoSchuh). The project Pay-per-Stress is financed with funding provided by the German Federal Ministry for Economic Affairs and Energy within the "Smarte Datenwirtschaft" program under project funding reference number 01MD19011. It is implemented by the DLR Project Management Agency.

References

[1] Rao, A.S., 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. PriceWaterhouseCooper. Accessed 13 April 2021.

[2] Wittpahl, V., 2019. Künstliche Intelligenz: Technologien| Anwendung| Gesellschaft. Springer Nature.

[3] Antonescu, D., 2018. Are business leaders prepared to handle the upcoming revolution in business artificial intelligence? Quality-access to Success 19 (166).

[4] 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. 30.06.2018 - 02.07.2018. ACM Press, New York, New York, USA, pp. 35–41.

[5] DIN 31051 Deutsches Institut für Normung e. V., 2012. Grundlagen der Instandhaltung. Beuth Verlag, Berlin.

[6] Hoffmann, F., Brockhaus, B., Metternich, J., Weigold, M., 2020. Predictive Maintenance für Schutzabdeckungen: Vom Geschäftsmodell zur Anwendung. WT Werkstattstechnik 110 (7-8), pp. 496–500.

[7] Serradilla, O., Zugasti, E., Zurutuza, U., 2020. Deep learning models for predictive maintenance: a survey, comparison, challenges and prospect. arXiv preprint arXiv:2010.03207.

[8] 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, pp. 191–199.

[9] Baptista, M., Sankararaman, S., Medeiros, I.P. de, Nascimento Jr, C., Prendinger, H., Henriques, E.M.P., 2018. Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. Computers & Industrial Engineering 115, pp. 41–53.

[10] Sharp, Michael, Thurston Sexton, and Michael P. Brundage, 2017. Toward semi-autonomous information. In: Lödding H., Riedel R., Thoben KD., von Cieminski G., Kiritsis D. (eds) Advances in Production Management

Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing. APMS 2017. IFIP Advances in Information and Communication Technology, vol 513. Springer, pp. 425–432.

[11] Benkedjouh, T., Medjaher, K., Zerhouni, N., Rechak, S., 2013. Remaining useful life estimation based on nonlinear feature reduction and support vector regression. Engineering Applications of Artificial Intelligence 26 (7), pp. 1751–1760.

[12] Engeler, M., Treyer, D., Zogg, D., Wegener, K., Kunz, A., 2016. Condition-based maintenance: model vs. Statistics a performance comparison. Procedia CIRP 57, pp. 253–258.

[13] Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2020. The M4 Competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting 36 (1), pp. 54–74.

[14] Chesbrough, H., Rosenbloom, R.S., 2002. The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. Industrial and corporate change 11 (3), pp. 529–555.

[15] Gassmann, O., Frankenberger, K., Csik, M., 2014. The business model navigator: 55 models that will revolutionise your business. Pearson, Harlow, England, London, New York, Boston, San Francisco, 387.

[16] Stähler, P., 2001. Geschäftsmodelle in der digitalen Ökonomie: Merkmale, Strategien und Auswirkungen. Josef Eul Verlag, Cologne.

[17] Chesbrough, H., 2007. Business model innovation: it's not just about technology anymore. Strategy & Leadership 35 (6), pp. 12–17.

[18] Reim, W., Åström, J., Eriksson, O., 2020. Implementation of Artificial Intelligence (AI): A Roadmap for Business Model Innovation. AI 1 (2), pp. 180–191.

[19] Schatsky, D., Muraskin, C., Gurumurthy, R., 2015. Cognitive Technologies: The real opportunities for business. https://www2.deloitte.com/us/en/insights/deloitte-review/issue-16/cognitive-technologies-business-applications.html/#endnote-5. Accessed 23 March 2021.

[20] Chesbrough, H., 2010. Business model innovation: opportunities and barriers. Long range planning 43 (2-3), pp. 354–363.

[21] Lee, J., Suh, T., Roy, D., Baucus, M., 2019. Emerging technology and business model innovation: the case of artificial intelligence. Journal of Open Innovation: Technology, Market, and Complexity 5 (3), 44.

[22] 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.

[23] Geissdoerfer, M., Savaget, P., Evans, S., 2017. The Cambridge Business Model Innovation Process. Procedia Manufacturing 8 (2), pp. 262–269.

[24] Osterwalder, A., Peigneur, Y., 2010. Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers, 1. Aufl. 2011 ed. Campus, Frankfurt.

[25] Osterwalder, A., Pigneur, Y., 2010. Business model generation: a handbook for visionaries, game changers, and challengers. John Wiley & Sons.

[26] Wagner, T., Tilly, R., Bodenbenner, P., Seltitz, A., Schoder, D., 2015. Geschäftsmodellinnovation in der Praxis: Ergebnisse einer Expertenbefragung zu Business Model Canvas und Co, in: Wirtschaftsinformatik, pp. 1298–1312.

[27] Hoffmann, F., Brockhaus, B., Metternich, J., Weigold, M., 2020. Predictive maintenance for protective covers: From business model to application. WT Werkstattstechnik (07-08), pp. 496–500.

[28] Stanula, P., Praetzas, C., Kohn, O., Metternich, J., Weigold, M., Buchwald, A., 2020. "Stress-oriented, data-based payment model for machine tools". Procedia CIRP 93, pp. 1526–1531.

[29] Bretones Cassoli, B., Ziegenbein, A., Metternich, J. Getting Started: KI zum Nutzen der Industrie vorantreiben. VDI-Z Integrierte Produktion, (E-Paper), Springer VDI, ISSN 0042-1766.

[30] Bruneel, J., D'Este, P., Salter, A., 2010. Investigating the factors that diminish the barriers to university–industry collaboration. Research Policy 39 (7), pp. 858–868.

[31] S. Ransbotham, S. Khodabandeh, R. Fehling, B. LaFountain, D. Kiron, 2019. Winning With AI. MIT Sloan Management Review and Boston Consulting Group. https://sloanreview.mit.edu/projects/winning-with-ai/.

[32] Andrew Ng, 2019. AI Transformation Playbook: How to lead your company into the AI era. Landing AI. https://landing.ai/ai-transformation-playbook/.

Biography

Beatriz Bretones Cassoli, 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. Her research interests include Knowledge Representation, Automated Planning and Predictive Maintenance.

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 interest are AI-driven Business Models in manufacturing and Predictive Maintenance.

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.

CPSL CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

A Process Model For Managing Business Applications In The Digital Transformation

Jan Hicking[1], Tobias Schröer[1], Florian Renneberg[1]
[1]Institute for Industrial Management, FIR at RWTH Aachen University, Campus-Boulevard 55, Aachen 52074, Germany

Abstract

The digital transformation brings up various new tasks to manage new business application software and integrate them into existing business processes and legacy systems, which are necessary to keep e.g. a production system running. Today, all these tasks are on the one hand not clearly defined and on the other hand, responsibility of these cross-disciplinary tasks is unclear in companies being mostly structured in a function-oriented way. While quality management has developed to a firmly established function of process excellence years ago, IT-application management is still to become an inevitable part of the digital transformation. There are just a few authors trying to define and describe this part, the related tasks, and necessary roles in an organization. In this paper, we show how the business needs of a company can influence the ideal adaptation of the digitization solutions and thus become the success of the digital transformation. We base the paper on a use case in manufacturing companies. We then describe how companies deal with business application systems today. Based on the framework Aachen Digital Architecture Management we describe how a company can holistically improve the management of business application systems.

Keywords

Business Applications; Digital Transformation; Digital Architecture Management; Process Management

1. Introduction

Digital transformation affects all companies in all industries and describes the influence of digital technologies on the change of value creation of companies [1]. The SARS-CoV-2 pandemic showed that processes can change in a flash when external influences exert so much pressure to change [2]. Manufacturing companies have not only felt this pressure since the pandemic; transformation was also important beforehand in order to remain competitive [3]. As early as 2016, many companies began to stimulate this change by introducing Chief Digital Officers (CDOs). This new role was expected to drive digital transformation from within the management board and drive it with commitment. To be sure, a CDO was capable of partially prototyping smart products and digital services and sometimes even launching them on the market. However, the CDO was only able to steer a complete digital transformation in exceptional strategic cases. The challenge lies in the fact that a CDO cannot control IT. This is because an IT manager is often assigned to a Chief Financial Officer (CFO). So, unlike digitalization, IT is not represented on the board. The obvious solution would be to appoint a Chief Information Officer (CIO) to the board. But this variant has not regularly led to the goal in the past either. Research and industry must understand that IT is more than just an IT topic [4]. Every company needs IT capabilities to run processes robustly and to test prototypes in different infrastructures. IT continues to be an enabler for innovation and efficiency. However, this should not be confused with the IT department. One of the most important digitization projects is the use of business applications such as Enterprise Resource Systems (ERP) systems not only for planning processes but also order tracking and tracing. Such systems are an inherent part of process execution so that there cannot be digital transformation without including an ERP system.

In this paper, we show how the business needs of a company can influence the ideal adaptation of the digitization solutions and thus become successful in digital transformation. We implement this introducing the use case management of business application systems in manufacturing companies. We first describe the initial situation in many manufacturing companies and the current handling of business application systems as well as the resulting challenge. Then, we present what support companies lack today based on frameworks that support the design of the architecture and, on the other hand, the IT capabilities of a company. We then introduce Aachen Digital Architecture Management as a structuring framework for digital transformation. Finally, we describe how a company can improve the management of business application systems holistic.

2. State of the art

2.1 The way manufacturing companies deal with business application software

Based on more than 100 projects on the topic of ERP system selection over the last 10 years, we have found that companies often only react instead of acting. Based on this empirical data, we have found that business application systems are replaced by new systems in regular cycles of 10 to 15 years. In most cases, this is because the performance of a system no longer meets the needs of a company. We consciously take into account that both the requirements and the performance of a system can vary. Exemplary for the external pressure of change induced by the pandemic is that a more tasks must be completed remotely, especially in administrative areas. Therefore, the system must be capable of allowing remote work. Technically, this can be solved in many ways. For example, through VPN access or a cloud application. The fact is that a company is constantly changing. Customers, partners and suppliers all have an influence. Companies follow the same process again and again: Unmet requirements create pressure on an existing system. Ultimately, they lead to the decision of implementing a totally new system. But even during the implementation phase of the new system, requirements have evolved again. However, they regularly do not influence the new system after initial completion of requirements engineering. This is exactly what happens afterwards in the next multi-year cycle based on our empirical data, see Figure 1. The benefit of a system is therefore only high in the short term and then constantly decreases over time.

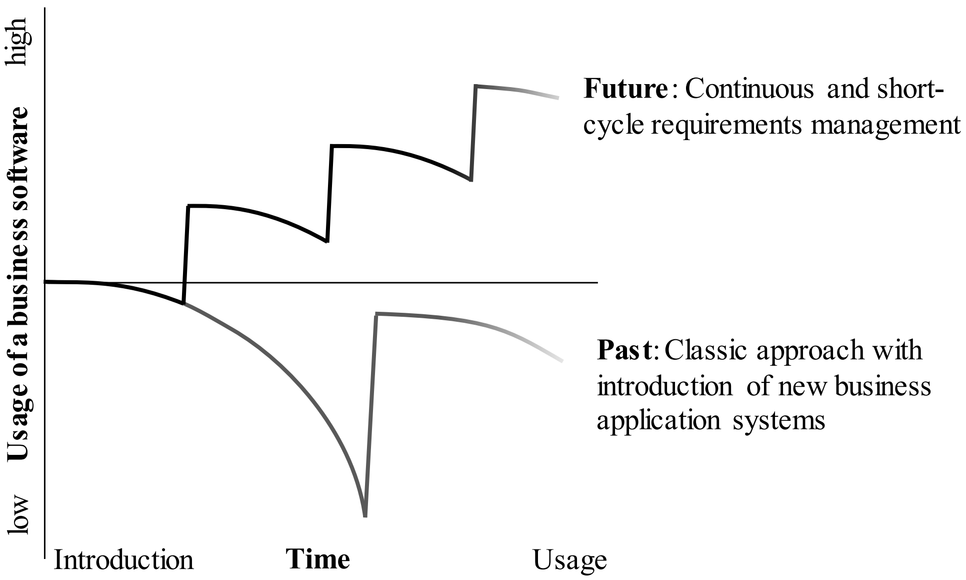


Figure 1: Managing business application systems in the digital transformation era

From this, the need for research can be derived that companies must learn that requirements must have an influence on the adaptation of systems. In companies, there is a strict separation between the procurement, implementation, maintenance and use of a business application system. Companies lack processual and organizational mechanisms for allowing requirements from the usage phase to flow into the further development of the system [3].

2.2 Existing frameworks of enterprise architecture and IT management

In this chapter we describe existing frameworks of EA and IT management and, by consolidating of IT capabilities and processes, show that there are only three relevant tasks to properly manage business application systems.

In general, the ITIL framework describes best practices in managing services [5]. The IT Infrastructure Library (ITIL) aims to form the optimal support of the business processes for the customers of an IT service provider with the help of information technology [6]. IT services are provided that add value for customers and fulfill both the utility (fit for purpose) and warranty (fit for use) characteristics. The core contents of the ITIL framework are the description of processes, functions, possible organizational structures and supplementary methods for implementation. This enables the IT service provider to achieve the stated goals. In order to map the entire life cycle of IT services, the necessary processes and functions are structured in 5 life cycle phases. [7] Companies of large and medium size use ITIL to improve the quality of services [5]. Next to numerous benefits that ITIL offers, some challenges can also be noted. For one, the documentation applications in ITIL are not mature. Likewise, only rough recommendations are given as to which processes should actually be implemented, so that no clear decision support is provided. [6] Other criticisms include the varying levels of detail in process descriptions, inconsistencies and contradictions between processes, barely applicable definitions of key performance indicators, and a lack of processes and procedures for managing processes and projects [7].

The Zachman Framework is a descriptive and holistic representation of an enterprise to gain insight into business processes [8]. The Zachman Framework is suitable for complex systems and incorporates numerous perspectives. These include the Executive (setting the agenda), the Business Management (who runs the organization), the Architecture (who identifies building blocks), the Engineer (who designs the building blocks), the Technician (who implements the database or workflow system), and the Enterprise (physical enterprise offices) as audience perspectives. [9] In the Zachman framework, engineered objects are defined according to the interrogatives *What? How? Where? Who? When?* and *Why?* in combination with the audience perspective [10]. Thus, the Zachman Framework can be described as an enterprise ontology, which has no tooling or process implication [8, 11].

COBIT is based on more than 40 detailed international IT standards, frameworks, guidelines and best practices. COBIT is located at the strategic level and integrates different standards and their regulatory objectives into a common framework. It establishes a link between corporate objectives and IT objectives at the core business and, to this end, provides metrics and maturity models to measure the achievement of objectives. COBIT is process-oriented, with 34 processes and four domains. [12] COBIT defines the domains as Planning and Organization, Acquisition and Implementation, Delivery, and Support and Monitoring [13]. In terms of content, COBIT covers the objectives of IT controls and measures that are intended to ensure the security of information in the company. For this purpose, criteria such as effectiveness, efficiency, confidentiality, integrity, availability, compliance, and reliability are applied to corporate information. [12]

TOGAF is a framework as well as a detailed methodology and provides various supporting tools for developing an enterprise architecture [14]. The standard is developed and maintained by The Open Group, an industry consortium. The framework assists in the acceptance, production, use, and maintenance of enterprise architecture. It is based on iterative process models supported by best practices and real-world sets of architectural assets [15]. The core content of TOGAF is the Architecture Development Method (ADM), which is used to develop architectures. This method includes creating the architecture framework, developing architecture content, transitioning, and governing the realization of architectures. The ADM is complemented by the Architecture Content Framework (ACF), which is structured according to the content of the metamodel. The metamodel provides single insights into the domains of TOGAF. Where TOGAF consists of the Business, Data, Application, and Technology Architecture domains [16].

All the previously described frameworks contain many processes that are described in detail. While ITIL is very specific in terms of IT processes, the Zachman framework is very generic for the design of a company's architecture. All frameworks provide insufficient support for companies to continuously adapt business applications in the utilization phase. We have condensed the IT Capabilities anchored in processes to three essential ones. We show the results of the entire analysis in Figure 2 and describe them below.

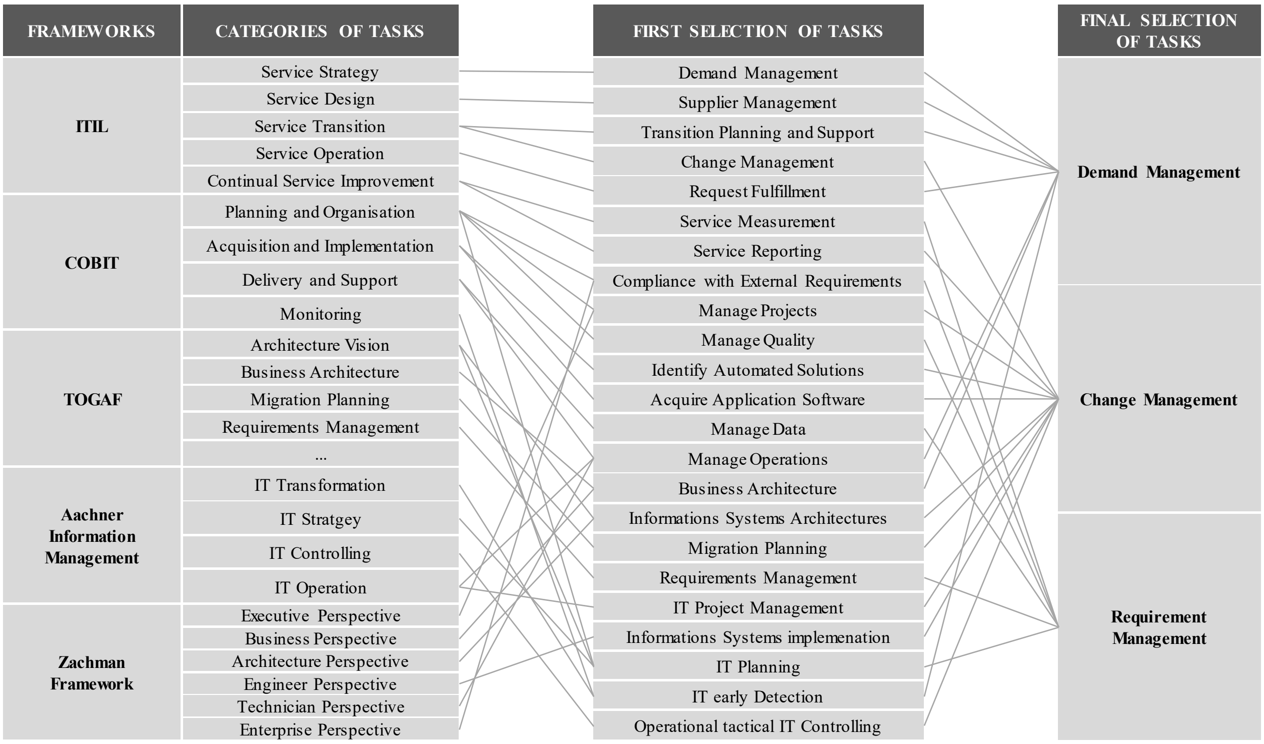


Figure 2: Consolidation of IT capabilities and processes

In the general context of production, demand management is the supply chain management process that balances the customer requirements with the capabilities of the supply chain [17]. In the context of this paper, demand management is understood as the identification of the need for digital solutions e.g., a new business application. The goal is the planning of an optimal range of digital solutions [7].

The goal of requirement management is a common understanding between the business and the IT department about the function to be supported and the framework conditions of the information structure as well as the joint evaluation of the requirements. This results in specific and systematically prioritized requirements that influences adoptions of existing digital solutions [18]. These comprise all the requests that are necessary to fulfil the objectives [14].

The aim of change management is to achieve the optimal design of the path from the starting point to the target point. The goal is thus to implement the optimal adaptation derived in the context of strategic management [19]. In the digital transformation context, change management encompasses the planning, release, coordination and acceptance of changes and additions to digital solutions e.g., business applications.

This is to ensure a clearly structured and controlled implementation of changes to digital solutions and the application landscape [7].

First, we showed that none of the existing frameworks cover the requirements of the digital transformation of a manufacturing company. Second, we described that there are only three relevant tasks to properly manage business application systems and that there is no need for a manufacturing company to dive deep existing and complex frameworks. However, there is a need for an overarching and strategic view of the management of business applications in the age of digital transformation, which we have already noted in the introduction. Therefore, we introduce the Aachen Digital Architecture Management (ADAM) in the following chapter.

3. Introduction of Aachen Digital Architecture Management

The Aachen Digital Architecture Management provides a framework that addresses the weaknesses of existing reference architectures while incorporating their strengths [3]. As a holistic model specifically developed for use by companies, ADAM structures the digital transformation journey of businesses in the areas of digital infrastructure and business development based on customer requirements. Companies are systematically empowered to drive forward the design of their digital architecture under consideration of various fields of action. Those responsible for digital transformation, from the top management team through to the operational innovation drivers in the specialist departments, need a systematic, structured approach that integrates the various digital transformation activities into a dynamic, scalable overall picture [3]. In this paper, ADAM represents the necessary foundation to introduce our model to manage business applications.

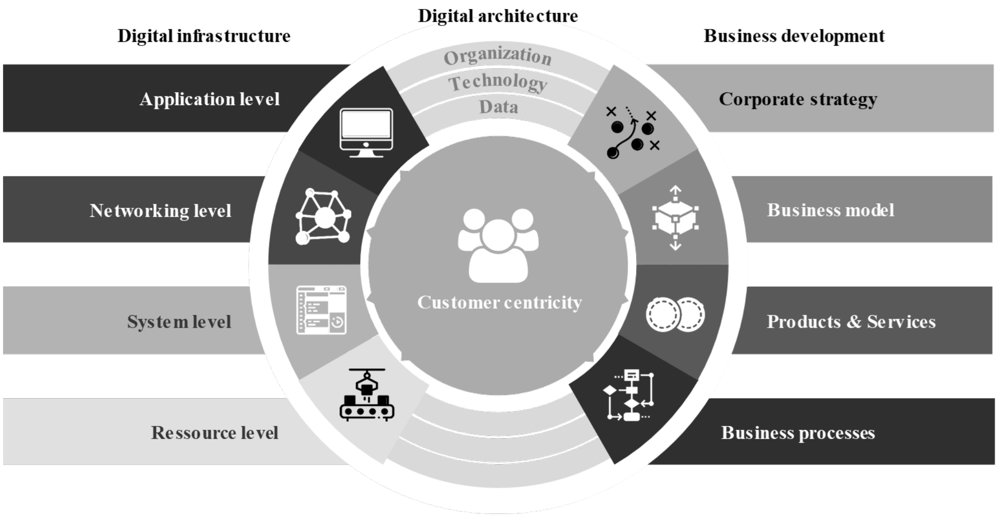


Figure 3: Aachen Digital Architecture Management

The Aachen Digital Architecture Management (Figure 3), considers two areas, digital infrastructure, which is subdivided into four design levels, and business development, which is subdivided into four development levels. The design and conceptualization of these levels, which is based on an analysis of internal and external customer requirements, provides the basis for building the digital architecture. The conceptualization of the four design levels of the digital infrastructure is based on frameworks established in research and practice for the description of interconnected companies, systems, and products. The design of the digital infrastructure enables a company to develop its business with the support of technology [3].

Customer: ADAM's focus is on the customer, whose needs and requirements are essential for the design of the digital architecture. It makes no difference whether it is an internal customer or a customer for physical or digital goods and services. Not only does the digital infrastructure have to be aligned with the specific business transformation fields, but also has to consider the requirements of customers who are to reap the benefits it provides. Corporate executives must take internal and external interest groups into account in all important decisions [3].

Digital Infrastructure: The structure and content of the digital infrastructure are defined on the basis of four design levels. The application level, refers to all user-centric applications that allow a user to use the corporate resource "information" in a simple and intuitive way as part of their value-creating activities. Due to rapidly changing circumstances, enterprise dashboards are subject to constant change. In order for these dashboards to be valid at all times and to be used effectively in managing the company, they require agile and flexible adaptability. The networking level enables a company to create interdependent, loose couplings of application and other data-providing systems. It ensures the availability of all relevant data and orchestrates the distribution of data between the different levels. Technologically essential components include IoT platforms and suitable communication technologies. Data virtualization is the key enabler, ensuring the user- and developer-friendly provision of data in the long term. The systems level contains static, slowly changing operational core systems e.g. ERP systems that form the information technology backbone of a company and manage and support central value-added processes. In addition to operational core systems, the systems level also contains other, unique data storage solutions, such as databases on the shop floor. Further, the systems level contains the business logic of a company. The resource level is comprised of production and production-related machines, equipment and other physical assets, employees, including their skills and competencies, intelligent products in the field, and software and hardware infrastructure for IT operations [3].

Business Development: The structure and content of business development are defined based on four development levels. Business development represents business activities in the well-known understanding of Business-IT alignment. The corporate strategy development level for business development determines the way in which the value of the company is increased in the long term. Based on clear corporate objectives, such as achieving a strategic position for success in the market, the company's corporate strategy, digitalization strategy, and IT strategy are formulated in a coordinated, integrated manner. The business models development level determines how companies act on the market to implement their corporate strategy. In all sectors, there is a particular focus on digital business models which make it possible for companies to successfully exploit the potential of digital transformation and open up new business segments. The Products & Services development level deals with the actual design of the value creation process to achieve real competitive advantage. The basis for this are service systems consisting of intelligent, interconnected products, smart services, and digital components. The business processes development level is concerned with the efficient design of business processes both internally within the company and with external process participants. For example, existing business processes must be optimized, designed end-to-end, and digitalized in order to scale digital transformation projects in the company [3].

Three architectural views provide analytical perspectives on the design levels: The organizational view, the technology view, and the data view offer a comprehensive picture of the four design levels of the digital infrastructure. The exhaustive analysis provided by the three views facilitates a comprehensive design of the digital infrastructure.

The organizational view provides the framework for the design of the digital architecture. From a conceptual point of view, the focus is on the development of a suitable management and control system and of the structural and process organization. The organizational view describes how staff interact with each other, with corporate systems, and with digital solutions. The organizational view also defines rights of use, responsibilities, and organizational affiliations. The technology view is characterized, in particular, by the selection of suitable technologies at the four design levels and their integration. Concept development places particular emphasis on deriving the required technologies from the organizational framework conditions, such as existing competencies. Technologies to be implemented include information and communication technologies, e.g. hardware, the operating environment, or the technical implementation of interfaces.

Requirements from the specialist areas and from business development, in particular, serve as input for the activities of the technology view. With the help of models across the different design levels, the data view offers a uniform perspective on the company's data. It describes the data structure, its components, and their interrelationships across the design levels. This ensures that the data and information required at the application level is available in sufficient quality and granularity and in the right structure. To this end, in the concept phase, a comprehensive information requirements analysis must be performed and specific data models down to the resource level must be developed. The data perspective deploys various tools to differentiate between terms like data and information and ensures the standardization of the data used in the company.

ADAM provides a framework for the most important elements in digital transformation. It serves as a framework for breaking down various issues into their components and structuring them. In the following chapter, we show how ADAM can be used to support the management of business application systems.

4. Process model of managing business applications within ADAM

In this chapter we describe how the ADAM framework can support the management of business application systems by embedding the three relevant tasks (Section 2.2) in ADAM.

ADAM first shows which elements are important in the digital transformation. Now we narrow down the consideration of our previously described use case to the system level (Figure 4, left-hand side) and the business development level (Figure 4, right-hand side). In order to deal adequately with business needs, a funnel is required that makes them recognizable in the organization. At this point, we emphasize that business needs represent internal needs on the one hand, but can also be driven by customer feedback and customer requirements.

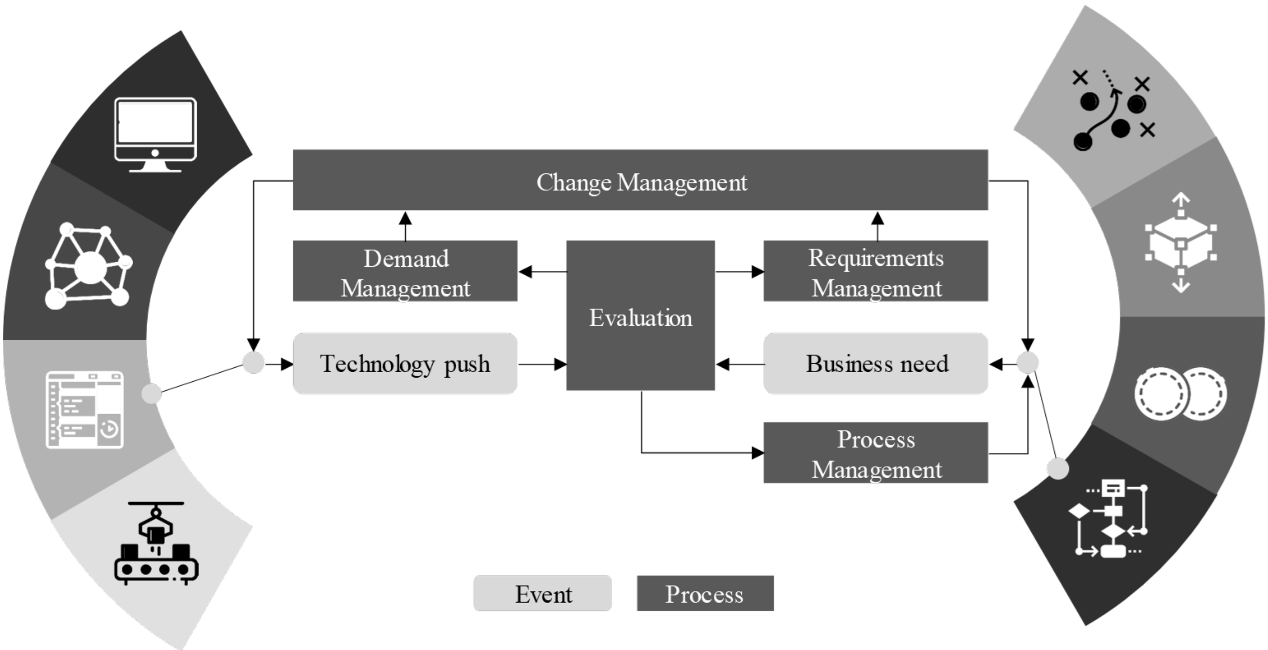


Figure 4: Process model for managing business applications within ADAM

We start by using an event as the input variable to the model. This is either a business need or a technology push. First, once a business need is communicated, it is evaluated by the "evaluation" committee. This committee comprises all relevant stakeholders of a company. There must be at least one representative from each design level and each business development level assigned to ADAM. This ensures that the interests of the entire company are taken into account. However, not only can business needs be communicated, but new technologies can also be introduced. We call this a technology push, and it means that a new technology is presented to the committee in order to serve as a source of ideas for all those involved. Business needs and technology pushes are regularly discussed in the "evaluation" committee. The committee jointly determines how to deal with the respective input. Three different follow-up processes can be initiated as a result of a committee decision.

First, the committee may conclude that the subject of discussion can be implemented with a simple process adjustment via the process management team. We hereby take up the aspect that not every problem has to be compensated with a technological solution. Process solutions can often be found to be even more effective because they address the root cause of a problem. The process management team receives all relevant information and gets in contact with the counterparts in the company who have reported a business need. Implementation is then initiated and carried out by the process management team in close exchange with the counterparts. If these process adaptions result in a new business need and potential IT demand, the team is encouraged to formulate a new request.

Second, the committee may conclude that the subject of discussion is a simple or complex adaptation of existing business application systems. Regardless of the difficulty and effort of the adaptation, the subject of discussion is passed on to “requirements management”. This works on a solution in accordance with ITIL, COBIT, etc. However, in direct exchange with the committee and the counterparts in the company who have reported the need. The committee also has the task of monitoring the status of an adaptation, whether there are challenges and how satisfied the counterparts are with the progress of the solution. The "evaluation" committee ensures that requests are not simply handed over without comment and either processed somehow or not at all.

Third, the committee may conclude that the subject of discussion is something completely new and unknown. A solution component that does not yet exist in the organization. In this case, the committee helps to hand over the discussion item to “demand management”. This works in accordance with ITIL, COBIT etc. on a solution procurement. However, in this case, too, in direct exchange with the counterparts in the company and the committee. The committee takes into account the fact that there are no existing solution modules in the company when making the decision. In this way, the committee prevents an ineffective accumulation of solution modules.

Finally, the results are continuously discussed with the “change management”. The “change management” has the task of ensuring that, on the one hand, partial solutions are fed back into the organization at an early stage. On the other hand, it must be ensured that the results of “demand and requirements management” are actually accepted by the organization. “Change Management” also evaluates the impact of the results. The entire model thus helps to ensure that the benefits of existing solutions are continuously improved. However, blind spots in the organization are also equipped with digital solutions.

5. Summary and Outlook

In this paper, we have shown how manufacturing companies can improve the management of their business applications by considering process, requirements and demand management in a structured manner. In doing so, we first used the results and findings of many software selection projects to illustrate what many companies regularly fail at. Particularly noteworthy is that due to a lack of communication and transparency, many solutions are introduced twice or are not available to all employees in an organization. Furthermore, we analysed the most common frameworks for enterprise architecture management and IT management. We found that all frameworks contain interesting components but are not effective in their entirety. Both analyses led to the research gap that there are no supporting models for enterprises. We then introduced the Aachen Digital Architecture Management framework. It sets out how the different components of digital transformation can be structured. The model served us to show the connection between technology and organization. Finally, we described the model to better manage business applications as well as new solutions. The core component is the "evaluation" committee, which helps to ensure that many relevant stakeholders on the side of the design levels and the business development levels jointly decide on important issues. In the future, this model will be used to develop a role model for the “evaluation” committee. This will identify the key digital capabilities in line with ADAM. It will provide companies with even more concrete support in driving forward the digital transformation effectively and transparently. Beside that, we use ADAM in strategic workshops with executives of manufacturing company to structure digital transformation topics and to show its practical applicability and relevance.

References

[1] Hess, T., Matt, C., Benlian, A. Options for Formulating a Digital Transformation Strategy. MIS Quarterly Executive 15 (2).

[2] Horgan, D., Hackett, J., Westphalen, C.B., Kalra, D., Richer, E., Romao, M., Andreu, A.L., Lal, J.A., Bernini, C., Tumiene, B., Boccia, S., Montserrat, A., 2020. Digitalisation and COVID-19: The Perfect Storm. Biomedicine hub 5 (3), 1341–1363.

[3] Hicking, J., Wenger, L., Abbas, M., Benning, J., Bremer, M., Clemens, F., 2020. Aachener Digital-Architecture-Management: Wegweiser zum digital vernetzten Unternehmen : Positionspapier. FIR e.V. an der RWTH Aachen, Aachen, 31 pp.

[4] Tabrizi, B., Lam, E., Girard, K., Irvin, V., 2019. Digital Transformation is not about technology. Change Management, Boston.

[5] O. G. B. Cabinet, 2011, cop. 2011. ITIL® service design, 2nd edition ed. TSO, London, 442 pp.

[6] Ahmad, N., Tarek Amer, N., Qutaifan, F., Alhilali, A., 2013. Technology adoption model and a road map to successful implementation of ITIL. Journal of Ent Info Management 26 (5), 553–576.

[7] Huber, M., Huber, G., 2011. Prozess- und Projektmanagement für ITIL®. Vieweg+Teubner Verlag, Wiesbaden.

[8] Gerber, A., Le Roux, P., Kearney, C., van der Merwe, A., 2020. The Zachman Framework for Enterprise Architecture: An Explanatory IS Theory, in: Hattingh, M., Matthee, M., Smuts, H., Pappas, I.O., Dwivedi, Y.K., Mäntymäki, M. (Eds.), Responsible design, implementation and use of Information and communication technology. 19th IFIP WG 6. 11 Conference on e-Business, e-Services, and e-Society, I3E 2020, Skukuza, South Africa, April 6-8, 2020, Proceedings, vol. 12066. Springer, Cham, pp. 383–396.

[9] Bondar, S., Hsu, J.C., Pfouga, A., Stjepandić, J., 2017. Agile Digital Transformation of Enterprise Architecture Models in Engineering Collaboration. Procedia Manufacturing 11, 1343–1350.

[10] Zachman, J.A., 1987. A framework for information systems architecture. IBM Syst. J. 26 (3), 276–292.

[11] Urbaczewski, L., Mrdalj, S., 2006. A comparison of enterprise architecture frameworks. IIS.

[12] Grünendahl, R.-T., Steinbacher, A.F., Will, P.H.L., 2017. COBIT und BSI als Leitschnur der IT-Sicherheit, in: Grünendahl, R.-T., Steinbacher, A.F., Will, P.H.L. (Eds.), Das IT-Gesetz: Compliance in der IT-Sicherheit. Leitfaden für ein Regelwerk zur IT-Sicherheit im Unternehmen, 3., korrigierte Auflage ed. Springer Vieweg, Wiesbaden, Germany, pp. 13–18.

[13] Astuti, H.M., Muqtadiroh, F.A., Tyas Darmaningrat, E.W., Putri, C.U., 2017. Risks Assessment of Information Technology Processes Based on COBIT 5 Framework: A Case Study of ITS Service Desk. Procedia Computer Science 124, 569–576.

[14] The Open Group Standard, 2018. The TOGAF® Standard, Version 9.2.

[15] Camatti, J.A., Rabelo, G.M., Borsato, M., Pellicciari, M., 2020. Comparative study of open IoT architectures with TOGAF for industry implementation. Procedia Manufacturing 51, 1132–1137.

[16] Mayer, N., Aubert, J., Grandry, E., Feltus, C., 2016. An Integrated Conceptual Model for Information System Security Risk Management and Enterprise Architecture Management Based on TOGAF, in: Horkoff, J., Jeusfeld, M., Persson, A. (Eds.), The Practice of Enterprise Modeling. 9th IFIP WG 8. 1. Working Conference, PoEM 2016, Skövde, Sweden, November 8-10, 2016, Proceedings, vol. 267. Springer, New York, pp. 353–361.

[17] Croxton, K.L., Lambert, D.M., García-Dastugue, S.J., Rogers, D.S., 2002. The Demand Management Process. The International Journal of Logistics Management 13 (2), 51–66.

[18] Deindl, M., 2013. Anwendungsorientiertes Informationsmanagement: Aufgaben und Methoden der Informationsmanagementforschung im RWTH Aachen Campus Cluster Logistik. Dissertationsvortrag, 2013.

[19] Lauer, T., 2019. Change Management. Springer Berlin Heidelberg, Berlin, Heidelberg.

Biography

Dr.-Ing. Jan Hicking (*1991) has been working at the Institute for Industrial Management (FIR) at the RWTH Aachen since 2016, first as a project manager and since 2020 as head of the department Information Management. He received his Ph.D. in 2020 in the field of smart products. In his research he focuses on the alignment of business processes and digital solutions.

Tobias Schröer, M. Sc. (*1991) has been working at the Institute for Industrial Management (FIR) at the RWTH Aachen since 2016, first as a project manager and since 2020 as head of the department Production Management. As such he leads a variety of applied research and consulting projects. In his research he focuses on the real-world applications of business software for operations in production and logistics.

Florian Renneberg B.Sc. (*1996) has been working for the division Information Management at FIR RWTH Aachen University since the beginning of 2021. As a research assistant, he supports various research and industrial projects in the field of digitalization. He holds a B.Sc. in Mechanical Engineering and Business Administration and is currently enrolled in the same M.Sc. program.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Analysing the state of digitisation in SME – A survey based on an SME-specific maturity model

Nicolas Wittine, Robin Sutherland, Sigrid Wenzel, Ana Luiza Amaral Bicalho
University of Kassel; Department of Organization of Production and Factory Planning, Germany

Abstract

The prevailing volatile changes in the market are forcing companies to perform increasingly complex planning tasks. Furthermore, shorter product life cycles and a more frequent adaption to customer requirements arise from a sellers' market shifting to a buyers' market. Regarding the digital factory planning, appropriate digital methods, tools, and models help master these new challenges. Depending on industry sectors and company size, the application and implementation of the methods and tools of the digital factory vary. Especially small and medium-sized enterprises (SMEs) show limited progress regarding digitisation due to a lack of expertise and qualified personnel. Thus, identifying suitable methods and tools for SMEs is essential for developing an implementation plan driving their digital transformation. Therefore, this article uses a survey analysing and classifying the situation of SME via an SME-specific maturity model. By investigating the correlations between the impacting variables, it is possible to identify the untapped potential, forming the basis for developing workshops and training to gain experience in dealing with methods and tools for digital factory planning.

Keywords

Digitisation; Digital Factory Planning; Organisation development; Industry 4.0; Maturity Index; SME

1. Introduction

The terms volatility, uncertainty, complexity, and ambiguity best describe the current industrial landscape [1]. In the wake of the ongoing corona pandemic, short response times and flexible change processes are vital abilities for businesses worldwide. Innovation, automation, and digitisation are more than ever of utmost importance in these times [2]. Over half of the European SMEs (see EU definition [3]) see reduced revenues and a fifth is even fearing to be unable to meet financial obligation [4]. Besides this out of ordinary pandemic, the digital change and a customer-driven switch to a buyers' market impact SMEs non the less [5, 6]. The digital factory offers a toolset dedicated to facilitating better planning capabilities and responsiveness [6]. To utilise these tools, companies need qualified and trained personnel. SMEs often lack those personnel and are overwhelmed by a holistic approach to digital change [6, 7]. The vast diversity of tools and methods suited for different business areas contributes to even more complexity within the workplace. Therefore, SMEs need to focus on the most critical aspects of the digital change and its primary drivers. This effort allows promoting change from within the company sustainably.

This paper aims to increase the understanding of SMEs' current situation regarding digitisation and examine its drivers. The core of this publication is an interview study based on an SME-specific maturity index. Using appropriate methods to analyse the data, the driving aspects behind successful digitisation – from a novice to a more advanced level – are identified and reflected regarding their related dimensions and indicators.

Therefore, the paper gives an overview of related works by detailing existing maturity models' content, shortcomings and why an SME-specific maturity model is necessary. The subsequent section explains the structure and procedure of the conducted interview-based survey. Consequently, an analysis of the obtained data allows for identifying SME-specific drivers for digitisation. A detailed description of the identified drivers offers valuable insight into connections and interrelationships of different fields inside the used maturity index. Finally, the paper offers a summary and points to future research.

2. Related work

The existing works regarding digital maturity assessment feature broad aspects regarding Industry 4.0 and offer companies a way for (self-)assessment. Since not all maturity models are also SME-specific or cover the same topics, reviewing existing literature is necessary. The maturity model used in this publication is part of a thesis at the Department of Production Organization and Factory Planning at the University of Kassel [8].

Based on more than 50 publications, the most cited and SME-oriented ones are used to build a comprehensive SME-specific maturity model. This model acts as a basis for the conducted survey and the following identification of drivers of digitisation. The models differentiate topics (dimensions) and sub-topics (indicators) with different levels to express the companies maturity. The overall structure, therefore, is an important metric to structure a comprehensive model.

Combining such an overview with specific classification criteria allows building an SME-specific maturity index. Since SMEs still struggle with implementing methods and technologies regarded as Industry 3.0 [9], the inclusion of a maturity level of zero, which does not require a complete implementation, is necessary. Furthermore, the model has to be comprehensible and straightforward enough to be accessible to a broad audience, acting as guidance and eliminating uncertainties. Especially identifying areas of action and concrete measures increasing the companies digitisation level is an issue for many SMEs. [10] Therefore the developed model has to contain transformation steps acting as direct guidance to increase the digitisation level. Also, SMEs are often family-owned and lead. The often tech-savvy and authoritarian leadership tend to struggle with scalability in a growing company [11], thus hindering effective and efficient decision making. Especially non-technical aspects like modern leadership principles or enabling collaboration and ownership thinking are considered essential in digital change [12].

The following table presents an overview of the different maturity models consisting of the number of dimensions, indicators and levels. Models claiming to be SME-specific or addressing mid-tier companies are marked accordingly with an "x"(see Table 1). Regarding the classification criteria, an assessment ranging from not partly to wholly fulfilled is available.

The inclusion of level zero rates not fulfilled when not existent or the maturity model starts with an already advanced stage. On the contrary, the wholly fulfilled rating describes a clear definition for a technical maturity below Industry 4.0. If the concept of an incomplete Industry 3.0 stage is existent, but incomplete definitions or missing indicators hinder the understanding, awards a partly fulfilled rating.

A comprehensible and straightforward model that follows an accessible concept containing descriptions and visual aids is rated wholly fulfilled. If a user needs additional consulting by external experts or critical content is left unexplained, the criterion is rated not fulfilled. An explanation and guidance sufficient for a more advanced user justify the rating partly fulfilled.

Since transformation steps are valuable guidance models containing explicitly depicted transformation steps achieve the rating of wholly fulfilled. No description, missing or unpublicised content regarding the transformation result in the not fulfilled rating. Logically, partly fulfilled describes models that leave out descriptions of levels and means of their improvement.

When an essential element of the maturity model consists of non-technical aspects like leadership, management or company culture, the rating wholly fulfilled is awarded. Vice versa, a missing non-technical dimension results in not fulfilled. Models only partly considering social or cultural aspects also only partly fulfil this criterion.

Table 1: Overview of existing maturity models [8]

Author	*Dimensions*	*Indicators*	*Levels*	*SME-specific*	*Level zero*	*comprehensible and straightforward*	*Transformation steps*	*Non-technical aspects*
Akdil et al. 2018 [13]	3	13	4		◑	◑	◑	◑
Anderl 2015 [14]	2	12	5		◑	●	●	○
Bibby, Dehe 2018 [15]	3	13	4		◑	◑	◑	◑
De Carolis et al. 2017 [16]	4	-	5		◑	◑	◑	◑
Geissbauer et al. 2016 [17]	7	-	4		○	◑	◑	◑
Gökalp et al. 2017 [18]	5	-	6		◑	○	○	○
Häberer et al. 2017 [19]	5	17	5	x	○	◑	○	◑
Jodlbauer, Schagerl 2016 [20]	3	25	10		◑	○	○	◑
Landwehr-Zloch, Eichfelder 2019 [21]	6	19	4	x	◑	◑	◑	◑
Leineweber et al. 2018 [22]	3	44	8	x	◑	◑	◑	◑
Leyh et al. 2016 [23]	4	-	5		○	◑	◑	○
Lichtblau et al. 2015 [24]	6	18	6		◑	◑	◑	◑
Rauch et al. 2020 [25]	5	42	5	x	◑	◑	●	◑
Sames 2021 [26]	5	27	5	x	◑	●	●	◑
Schuh et al. 2017 [27]	4	31	6		●	◑	◑	●
Schumacher et al. 2016 [28]	9	62	5		◑	◑	◑	◑
Trotta, Garengo 2019 [29]	5	11	5	x	◑	◑	○	◑
State of fullfilement:				not ○		partly ◑		wholly ●

Many of the reviewed models have specific areas they excel in, but no single one is suited for a comprehensible, SME-specific maturity model. Yet, regarding the individual assessments, the entries in bold are considered as particularly appropriate to develop a comprehensive model. The derived SME-specific maturity model from the analysed literature, as depicted in Figure 1, features 6 different dimensions and 26 underlying indicators. Beginning with a level 0, the model details 5 levels, each with a description of its fulfilment requirements. The dimension “Production”, for example, contains information and communications technology (ICT) or monitoring, quality management (QM) & maintenance. Whereas “leadership, management and culture” deals with corporate and failure culture. The latter divides into levels beginning with a working environment shaped by blame (level 0) and finishing with a public analysis of failure (level 4).

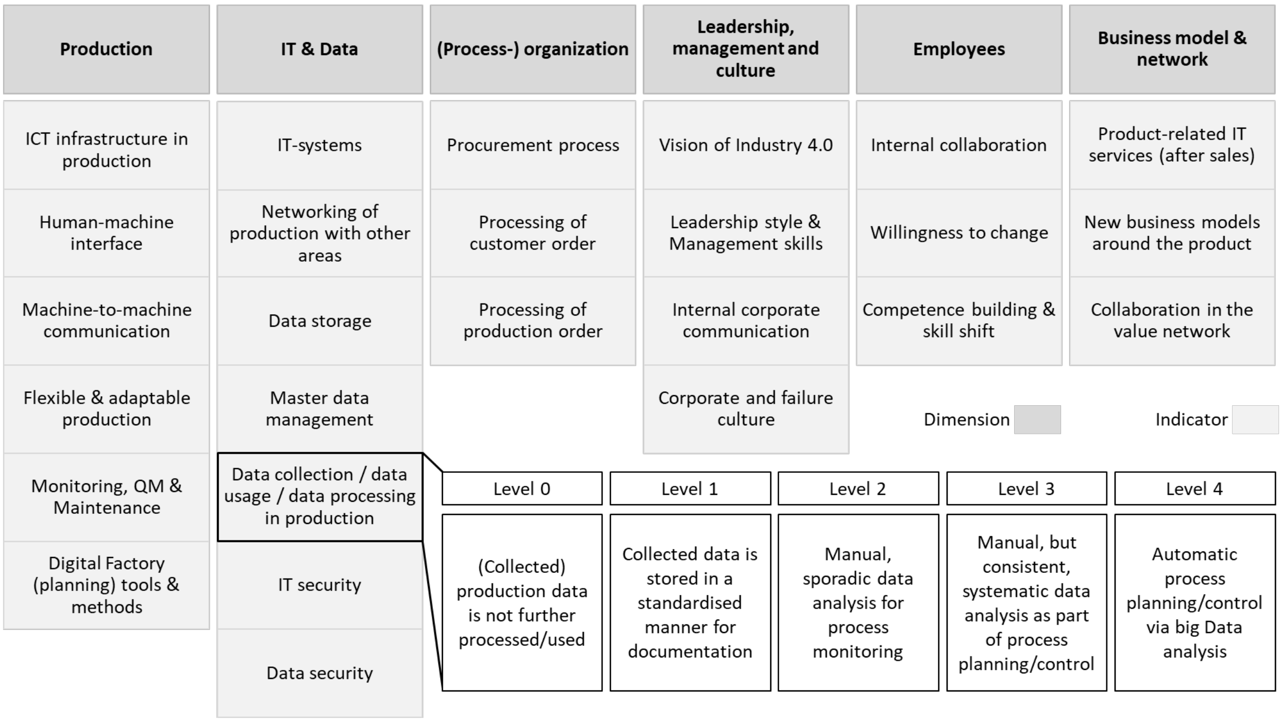


Figure 1: The SME-specific Industry 4.0 maturity model

3. Interview framework

An empirical study aims to grade SMEs' situation regarding their digital maturity based on the developed maturity model. The study uses verbal interviews of company representatives and their experts from different fields. The questions asked explicitly target the individual indicators in a sequence from the lowest to the highest maturity level; in addition to a mere classification, the subsequent analysis uncovers the potential for improvement by identifying maturity clusters, correlations and deriving needs for action in practice.

The participants in the survey are from nine different branches, all originating in the manufacturing sector SMEs from the region of Hesse in the middle of Germany. Interview partners are employees in crucial positions, including managing directors, production, process, IT or digitisation managers. The sample comprises 11 companies from different branches representing one small, eight medium-sized and two large enterprises (acc. to EU definition [3]). However, the latter fall under the more broad mid-tier concept [30], making their evaluation results still relevant. The validated interview guideline used is the outcome of a master thesis created at the research institute. The interview takes about 4 hours and is conducted either at the companies' sites or within a web meeting. The high complexity of the interview and its long duration tend to cause respondent fatigue. Therefore, the interview guides through the different topics in descending order of complexity, counteracting the loss of attention throughout the interview. In total, the guideline contains 32 guiding questions derived from the indicators and over 100 sub-questions detailing the explicit level. E.g. regarding the data collection / data usage / data processing in production (see Figure 1), the leading question of how production data is handled starts the dialogue. Follow up questions concerning whether, how and for what purpose the collection and analysis of production-related data deepen the topic. Also, general information, such as annual sales or the number of employees, are requested. For visual and thematic guidance, an accompanying presentation contains the current dimension of the questions asked. A transcript writer prepares a protocol for documentation purposes and later acts as a data basis for the classification in the SME-specific maturity model. Additionally, recording the audio is highly advised but needs the interviewees' consent beforehand.

4. Analysis of the interview

Each consecutive interview extends the database containing each companies maturity profile and overall maturity rating. The profile offers valuable insights into every companies area for improvement and targets specific topics in particular. While this information is valuable for each participant, it also enables analysis potential for a broader view regarding the state of maturity regarding SME. Since the database consists of only a small sample size regarding the number of questions asked, not every data analysis method is suited. A worthwhile goal is to classify the companies and to check for overarching similarities and differences. Another meaningful insight is the identification of drivers for digitisation.

The cluster analysis offers means to identify unknown groups (clusters) inside a given set of data and is viable to explore the conducted interviews. The calculated silhouette coefficient (S_c) classifies the clusters and represents how well elements can be assigned to a specific cluster, dividing into weak ($0.25 < S_c \leq 0.50$), medium ($0.50 < S_c \leq 0.75$) and strong structures ($0.75 < S_c \leq 1.00$). [31] In the present case, the analysis finds two distinct clusters ($S_c = 0.56$) across all dimensions (see Figure 2). The primary differentiation seen in the spider chart breaks down to Cluster 1, containing beginner companies and cluster 2 advanced companies. The more advanced dimensions are "Leadership, management and culture", "employees", "IT & Data", and "Business model & network". Therefore, those dimensions seem to be essential to advance a companies maturity level. Organization and production are dimensions that lag behind the overall in the more advanced companies and also highlight a need to catch up.

Additionally, the found clusters further prove the validity of the underlying maturity model. The found groups are distinct and have no overlap coinciding with the overall maturity level. Each company in cluster 1 also scores an overall maturity level of 1. Whereas the companies belonging to cluster 2 achieve level 2. The levels 0, 3, and 4 are missing in the surveys data set, and as a result, a corresponding cluster is also nonexistent, illustrate the current state of digitisation in the sampled SME.

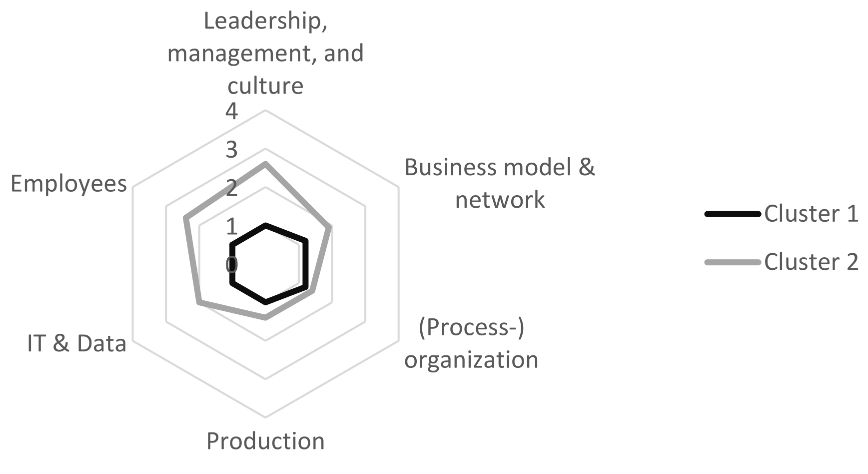


Figure 2 Result of the cluster analysis

The clustering already enables differentiated recommendations for various levels of digitisation and allows for a further investigation of specific drivers. The correlation analysis offers insight into relationships between the different dimensions and the overall maturity level and quantifies the direction and strength of a found correlation [32]. A correlation coefficient of +1 means that the two variables are positively related, whereas a value of -1 indicates the opposite. The small and not normally distributed dataset used in the paper requires a particular correlation analysis method, the Kendall tau (τ) correlation, and prevents false associations due to low significance [33].

Table 2 shows the correlation analysis results emphasising the correlations between the individual dimensions and the overall maturity level. Also, the table contains the significance of the found values. The relevance of a dimension derives from the respective τ-value and its effect strength from the *r*-value (weak $0.1 < r \leq 0.3$, medium $0.3 < r \leq 0.5$ and strong $0.5 < r$) [34]. The table contains the effect strength only for relevant values ($p < 0.05$), highlighting three dimensions, "IT & data", "Management, leadership and culture", and "employees". Those dimensions show a strong correlation with the overall maturity level and are also strongly related to other dimensions.

Table 2: Results of the correlation analysis

τ *r*	*Production*	*Organisation*	*IT & Data*	*Management, leadership and culture*	*Employees*	*Business model and network*	*Overall maturity*
Production							
Organisation	.195 -						
IT & Data	.242 -	.086 -					
Management, leadership and culture	.109 -	.101 -	.617 .845				
Employees	.160 -	.147 -	.617 .845	.487 .801			
Business model and network	.175 -	.078 -	.380 .639	.359 .651	.336 .622		
Overall maturity	.242 -	.086 -	.798 .953	.617 .845	.617 .845	.380 .639	

Regarding the identified drivers for digitisation, the cause, causation, and interconnectivity need consideration. The following statements can be formulated by further examining the identified dimensions and summarises the found relationships forming a concise overview of the identified drivers.

Specific IT infrastructures must be available to implement the digital transformation to Industry 4.0, enabling components or systems inside the factory and interconnecting the value-adding network. In addition to the information technology infrastructure, aspects of the collection, storage, and processing of data, including its use and maintenance. These aspects are necessary for primarily data-based decision making thus supporting the management. A digitally transformed enterprise requires consistent data in real-time to implement the digital factory's methods and tools, especially IT and data security. For example, ICT must be in place so employees, products, and machines can interact. ICT is essential for implementing digital factory methods and cross-divisional and cross-company collaboration and enabling new business models. The correlation analysis highlights IT infrastructure's influence regarding the overall maturity level, appointing this dimension a driver of digitisation. A low correlation with production is identifiable, but the effect strength is negligible. Even though the overall maturity benefits significantly from a strong IT & Data dimension, it acts as a supporting role, enabling, supporting and connecting the other dimensions.

Besides the companies ICT, the leadership has a strong influence on the overall maturity. Especially a well-defined Industry 4.0 vision and strategy is as essential as familiarity with modern management concepts. It plays a decisive role in motivating the employees to work together and achieve a common goal. Further, the management needs the will to drive and implement changes and provide suitable communication channels for internal networking. Strong support of the employees enables another major influencer of the overall maturity. However, in addition to structural changes, management must prepare employees for the digital factory's new requirements. The increasingly higher IT competencies lead to shifts in qualifications, which management must counteract with suitable training programs. To be prepared for the general changes in addition to the technical requirements, an open and innovative corporate culture must be in place and practised by management. An open failure culture that treats failures as means to improve and evolve existing technologies and processes. Such cultural aspects can promote new forms of internal collaboration making all those aspects essential for a successful digital transformation.

Even though employees strongly correlate with the overall maturity, they need enabling from management as leverage. This dimension strongly correlates with management, leadership and culture since employees form the connection between essential elements in business, highlighting the need for a non-technical dimension. Also, communication and collaboration between employees benefit highly from an IT and data-driven surrounding, making this interconnection logical. Nevertheless, the employees inside a company cannot drive the digital change without the digital framework and the support from the management, therefore even though essential employees are not considered direct drivers.

In conclusion, the state of digitisation in SME derives by combining the found drivers with the identified clusters. Especially the lack of higher scoring SME indicates their necessity for support and further research. Whether improving the drivers also acts as guidance for digitisation needs evaluation and validation.

5. Summary and outlook

In this paper, an interview study is evaluated based on an SME-specific maturity model. The model and the analysis results are explained and increase the understanding of the current situation of SMEs regarding digitisation. A performed cluster analysis classifies the different companies in the sample depending on their digitisation progress and validates the used maturity model. Additionally, further analysis has made it possible to identify and describe two digitisation drivers in the context of the used maturity model. Thus, SMEs should focus on the dimensions “IT & Data” and “Management, Leadership and Culture” in the early phases. Management that promotes employees and leads with a vision forms the foundation for digital change. Also, the availability of data and the resulting transparency enable informed and efficient decision making. This insight enables SMEs to focus on only a few key areas facing the constantly changing market’s challenges. Since the surveyed companies only ranked level 1 or 2, a larger data set containing more advanced companies (level 3 or 4) should confirm these findings and whether the interconnections and drivers change due to digitisation progress. Also, the reason for the lacking involvement of the production and organisation needs further investigation. Ultimately, based on the identified drivers, specific recommendations for SMEs can now be formulated. Besides practical guidance, future research should contain the necessary competencies and suitable ways to acquire them.

Acknowledgements

The research project – SME Knowledge Transfer Laboratory “Digital Factory” - (KMU-Wissenstransferlabor “Digitale Fabrik“) with a duration of three years (09/2019 - 08/2022) is funded by the European Regional Development Fund (ERDF - or EFRE in German) 2014 to 2020 (IWB-ERDF Programme Hesse) and by the State of Hesse.

References

[1] A. Priyono, A. Moin and V. N. A. O. Putri, “Identifying Digital Transformation Paths in the Business Model of SMEs during the COVID-19 Pandemic,” Journal of Open Innovation: Technology, Market, and Complexity, vol. 6, no. 4, 2020.

[2] H. Guo, Z. Yang, R. Huang and A. Guo, "The digitalization and public crisis responses of small and medium enterprises: Implications from a COVID-19 survey," Front. Bus. Res. China (Frontiers of Business Research in China), vol. 14, no. 1, pp. 14-19, 2020.

[3] European Union, "COMMISSION RECOMMENDATIONof 6 May 2003 concerning the definition of micro, small and medium-sized enterprise," Official Journal of the European Union, no. 124, pp. 36-41, 2003.

[4] J. Dimson, Z. Mladenov, R. Sharma and K. Tadjeddine, "McKinsey & Company," October 2020. [Online]. Available: http://bit.ly/mckinsey20. [Accessed 21 04 2021].

[5] U. Sendler, Industrie 4.0 grenzenlos, Berlin; Heidelberg: Springer Vieweg, 2016.

[6] U. Bracht, D. Geckler and S. Wenzel, Digitale Fabrik - Methoden und Praxisbeispiele, Berlin: Springer Vieweg, 2018.

[7] C. Faller and D. Feldmüller, "Industry 4.0 Learning Factory for regional SMEs," Procedia CIRP, vol. 32, p. 88–91, 2015.

[8] K. A. Schneider, "Entwicklung einer KMU-spezifischen Einschätzungsmethode der Unternehmenssituation im Kontext von Industrie 4.0.," Kassel, 2020.

[9] S. Mittal, M. Ahmad Khana, D. Romero and T. Wuest, "A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs)," Journal of Manufacturing Systems, vol. 49, pp. 194-214, 2018.

[10] S. Wiesner, P. Gaiardelli, N. Gritti and G. Oberti, "Maturity Models for Digitalization in Manufacturing - Applicability for SMEs," in Advances in Production Management Systems. Smart Manufacturing for Industry 4.0, Cham, Springer International Publishing, 2018, pp. 81-88.

[11] J. Bischoff, "Erschließen der Potenziale der Anwendung von Industrie 4.0 im Mittelstand. Kurzfassung der Studie," agriplan GmbH, Mühlheim an der Ruhr, 2015.

[12] H. Reinemann, Mittelstandsmanagement: Einführung in Theorie und Praxis, Wiesbaden: Springer Fachmedien, 2019.

[13] K. Y. Akdil, A. Ustundag and E. Cevikcan, "Maturity and Readiness Model for Industry 4.0 Strategy," in Industry 4.0: Managing The Digital Transformation, Cham, Springer International Publishing, 2018, p. 61–94.

[14] R. Anderl, "Leitfaden Industrie 4.0: Orientierungshilfe zur Einführung in den Mittelstand," VDMA Verlag, Frankfurt am Main, 2015.

[15] L. Bibby and B. Dehe, "Defining and assessing industry 4.0 maturity levels – case of the defence sector," Production Planning & Control, vol. 29, no. 12, p. 1030–1043, 2018.

[16] A. d. Carolis, M. Macchi, E. Negri and S. Terzi, "A Maturity Model for Assessing the Digital Readiness of Manufacturing Companies," Advances in Production Management Systems. The Path to Intelligent, Collaborative and Sustainable Manufacturing. APMS 2017. IFIP Advances in Information and Communication Technology, vol. 513, pp. 13-20, 2017.

[17] R. Geissbauer, J. Vedso and S. Schrauf, "Industry 4.0: Building the digital enterprise," 2016. [Online]. Available: http://bit.ly/geissbauer2016. [Accessed 21 04 2021].

[18] E. Gökalp, U. Şener and P. E. Eren, "Development of an Assessment Model for Industry 4.0: Industry 4.0-MM," Software Process Improvement and Capability Determination. SPICE 2017. Communications in Computer and Information Science, vol. 770, p. 128–142, 2017.

[19] S. Häberer , L. K. Lau and F. Behrendt, "Development of an Industrie 4.0 Maturity Index for Small and Medium-Sized Enterprises," IESM 2017 - 7th International Conference on Industrial Engineering and Systems, pp. 129-134, 2017.

[20] H. Jodlbauer and M. Schagerl, "Reifegradmodell Industrie 4.0. Ein Vorgehensmodell zur Identifikation von Industrie 4.0 Potentialen," INFORMATIK 2016, Gesellschaft für Informatik e.V., p. 1473–1487, 2016.

[21] S. Landwehr-Zloch and M. Eichfelder, "Checkliste zur Eigenbewertung der Industrie 4.0 - Readiness für kleine und mittlere Unternehmen (KMU)," Controller Magazin, vol. 44, no. 4, pp. 74-76, 2019.

[22] S. Leineweber, T. Wienbruch, D. Lins, D. Kreimeier and B. Kuhlenkötter, "Concept for an evolutionary maturity based Industrie 4.0 migration model," 51st CIRP Conference on Manufacturing Systems, vol. 72, pp. 404-409, 2018.

[23] C. Leyh, T. Schäffer, K. Bley and S. Forstenhäusler, "Assessing the IT and Software Landscapes of Industry 4.0-Enterprises: The Maturity Model SIMMI 4.0," Information Technology for Management: New Ideas and Real Solutions. ISM 2016, AITM 2016, vol. 277, p. 103–119, 2017.

[24] K. Lichtblau, V. Stich, R. Bertenrath, M. Blum, M. Bleider, A. Millack, K. Schmitt, E. Schmitz and M. Schröter, Industrie 4.0-Readiness, Aachen, Köln: IMPULS-Stiftung des VDMA, 2015.

[25] E. Rauch, M. Unterhofer, R. A. Rojas, L. Gualtieri, M. Woschank and D. T. Matt, "A Maturity Level-Based Assessment Tool to Enhance the Implementation of Industry 4.0 in Small and Medium-Sized," Sustainability, vol. 12, no. 9, pp. 1-18, 2020.

[26] G. Sames, "Reifegradmodell zur Digitalisierung und Industrie 4.0," THM-Hochschulschriften Band 18, Gießen, 2021.

[27] G. Schuh, R. Anderl, J. Gausemeier, M. t. Hompel and W. Wahlster, Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten, München: Herbert Utz Verlag, 2017.

[28] A. Schumacher, S. Erol and W. Sihn, "A Maturity Model for Assessing Industry 4.0 Readiness and Maturity of Manufacturing Enterprises," Procedia CIRP, vol. 52, pp. 161-166, 2016.

[29] D. Trotta and P. Garengo, "Assessing Industry 4.0 Maturity: An Essential Scale for SMEs," Proceedings of 2019 8th International Conference on Industrial Technology and Management (ICITM), pp. 69-74, 2019.

[30] B. Günterberg and G. Kayser, "SMEs in Germany - Facts and figures 2004," Institut für Mittelstandsforschung (IfM) Bonn - Materialien, Bonn, 2004.

[31] M. Ester and J. Sander, Knowledge discovery in databases: Technicken und Anwendungen, Springer-Verlag, 2013.

[32] K. Backhaus, B. Erichson, W. Plinke and R. Weiber, Multivariate Analysemethoden - Eine anwendungsorientierte Einführung, Berlin: Springer Gabler, 2018.

[33] S. Arndt, C. Turvey and N. C. Andreasen, "Correlating and predicting psychiatric symptom ratings: Spearmans r versus Kendalls tau correlation," Journal of Psychiatric Research, vol. 33, no. 2, p. 97–104, 1999.

[34] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, New York: Academic Press, 1988.

Biography

Sigrid Wenzel is a professor and head of the Department of Organization of Production and Factory Planning, University of Kassel. In addition to this, she is a board director of the Arbeitsgemeinschaft Simulation (ASIM), spokesperson for the ASIM Section Simulation in Production and Logistics, member of the advisory board of the Association of German Engineers Society of Production and Logistics (VDI-GPL), and head of the Committee Modelling and Simulation of the VDI-GPL.

Nicolas Wittine has been a research assistant at the Department of Organization of Production and Factory Planning at the University of Kassel since 2019. After obtaining his Master of Science in 2016, he worked as a development engineer in the automotive industry. He gained first-hand experience in developing safety-critical components, project management and industrialisation. Currently, his work focuses on traceability and digitisation.

Robin Sutherland has been a research assistant at the Department of Organization of Production and Factory Planning at the University of Kassel since 2020. He completed his master's degree in 2020 at the Karlsruhe Institute of Technology. Currently, his work focuses on Industry 4.0 and factory planning.

Ana Luiza Amaral Bicalho has been an industrial engineering master student at the University of Kassel since 2018. She has a bachelor's degree in production engineering focusing on simulation from the University of Brasília. Currently, she works as a student assistant at the Department of Organization of Production and Factory Planning at the University of Kassel.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Integrating Energy Flexibility in Production Planning and Control - An Energy Flexibility Data Model-Based Approach

Lukas Bank[1], Simon Wenninger[2], Jana Köberlein[1], Martin Lindner[3], Can Kaymakci[4], Matthias Weigold[3], Alexander Sauer[4] and Johannes Schilp[1]

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

Abstract

Production companies face the challenge of reducing energy costs and carbon emissions while achieving the logistical objectives at the same time. Active management of electricity demand, also known as Demand Side Management (DSM) or Energy Flexibility (EF), has been recognized as an effective approach to minimize energy procurement costs for example by reducing peak loads. Additionally, it helps to integrate (self-generated, volatile) renewable energies to reduce carbon emissions and has the ability to stabilize the power grid, if the incentives are set appropriately. Although production companies possess great potential for EF, implementation is not yet common. Approaches to practical implementation for integrating energy flexibility into production planning and control (PPC) to dynamically adapt the consumption to the electricity supply are scarce to non-existent due to the high complexity of such approaches. Therefore, this paper presents an approach to integrate EF into PPC. Based on the energy-oriented PPC, the approach identifies and models EF of processes in a generic energy flexibility data model (EFDM) which is subsequently integrated in the energy-oriented production plan and further optimised on the market side. An application-oriented use case in the chemical industry is presented to evaluate the approach. The implementation of the approach shows that EF can have a variety of characteristics in production systems and a clear, structured, and applicable method can help companies to an automated EF. Finally, based on the results of the use case, it is recommended to introduce EF in production companies stepwise by extending existing planning and scheduling systems with the presented approach to achieve a realization of flexibility measures and a reduction of energy costs.

Keywords

Demand Side Management (DSM); Energy Flexibility (EF); Production Planning and Control (PPC); Energy-Flexible PPC; Energy Flexibility Data Model (EFDM)

1. Introduction

In order to achieve the goals of the international climate agreement, the German government has decided to phase out coal and nuclear power generation [1,2]. Meanwhile, the share of electricity generation from wind and solar energy has to increase to achieve greenhouse gas neutrality of the electricity generated and consumed in Germany by 2050 [3]. As power generation from renewable energy sources such as wind and solar energy is limited in its ability to adjust to current energy demand, the associated fluctuating power generation poses major challenges for power grids and a significant increase in electricity price volatility is to be expected [5]. One possible solution to address these challenges is the increase in industrial Demand Side Management (DSM) also known as energy flexibility (EF). The concept of the energy-flexible factory takes advantage of being able to adapt quickly to price changes in the energy or balancing power markets with very little financial effort, thus derive an economic benefit [4,5]. Production planning and control (PPC) plays a central role in the context of the energy-flexible factory [6], as production is the main power consumer of a factory [7]. A large number of publications have dealt with energy-oriented PPC, in which production is optimised according to electricity price forecasts with regard to minimum electricity costs [8]. The PPC includes production planning, which plans the operations in long to medium term, and production control, which releases and controls the orders based on the previous planning [9]. Due to unavoidable disturbances such as machine breakdowns, a complete realization of the production plan is almost impossible. Therefore, the task of production control is to implement the production plan in the best possible way. The focus of PPC is to achieve logistical objectives, such as the timely completion of customer orders or machine capacity utilization [9]. With increasing awareness of sustainable aspects and pressure to optimise costs, additional objectives such as minimizing electricity costs or avoiding peak loads have been considered in PPC under the term energy-oriented PPC in recent years [10]. The approaches used to the present are either very rigid because they are based on forecasts and thus cannot react flexibly to short-term changes in different energy markets or are very complex.

EF must be considered without jeopardizing logistical objectives. In order to be able to react to changing energy prices at a later point in time and to be able to use EF, an approach is pursued in which EF measures are already taken into account in the planning phase, so that they can be taken when required. The approach thus offers two advantages. First, a company can adapt to a changing energy market without having to make unplanned changes to its production schedule. On the other hand, companies without energy-oriented planning can use this simple approach to at least consider EF measures and thus participate in the market. In doing so, the approach can be integrated into existing PPC methods, as they can be scheduled as additional orders, so to speak. Furthermore, by using an energy flexibility data model (EFDM), a description of the flexibility is used that is also suitable for marketing the flexibility, thus creating a continuous flow of information from production to the energy market. As publications usually stem from production or energy related domains, the interface between these domains has been insufficiently considered so far. In addition to the possibility of a reactive influence in the production control, due to market signals, this procedure also enables trading with EF for example via Intraday, Day-Ahead or the balance markets. Our approach is, therefore, to be understood explicitly as an extension of existing PPC solutions.

In addition to reviewing previous work and describing our own methodological concept, a use case from the German chemical industry is considered in order to validate and exemplify the application of the approach.

2. State of the Art

There is a variety of research in the domain of energy-oriented PPC and ongoing digitalization in manufacturing companies [11–16]. These works are usually based on electricity price forecasts with the goal of minimizing electricity cost due to production activities. Heinzl [11] and Wang et al. [16], as representatives for considering energy costs in production planning, use optimisation methods to combine the logistic and energy objectives. Süße et al. [15] additionally consider a battery storage system to account for the energy demand of production by means of a suitable battery charging management system. Even though the electricity load profile is altered in [15], energy flexibilities other than batteries are not yet considered.

In production control, they may react to short-term changes in electricity prices or due to changes in expected power consumption. Or they may react when the actual electricity price deviates significantly from the electricity price forecast [17,14]. In Schulz et al. [18] a closed-loop control was developed which is based on an algorithm with several data inputs such as production process data, environmental and energy data. The goal was the minimization of residual loads in the factory, but market interactions were not taken into account. Summarizing, none of the previous publications enable for flexibility trading.

Previous approaches do not take the information requirements of energy markets into account because, as mentioned, they treat energy prices as a given input variable. However, a continuous view from production to energy market requires the possibility of a uniform and simple description of energy flexibility that is still accurate. Different energy sources, dependencies between the production infrastructure and the supply technology, and restrictions due to logistical objectives as well as the involvement of different IT systems are just some of the general conditions that have to be considered when dealing with EF in production [19]. An energy flexibility model which addresses all these points is published by Schott et al. [20]. EFDM provides the framework for generically describing flexibilities and flexibility measures using key figures and technical parameters. A flexibility describes the potential possibilities of an energy flexible system to vary its performance compared to the reference operation. EFDM uses the classes *flexible load*, *dependency, storage,* and *flexible load measure*. The degrees of freedom of the EF – respective the flexibility potential – are defined by the characteristics of the key figures of the mentioned classes. A flexible load models a technical system or the interaction of different technical systems that have the potential to evoke a change in energetic performance. The *dependency class* is used to model constraints and dependencies for the interaction of multiple flexible loads. The use of a flexible load can, for example, imply or exclude the use of another flexible load. A storage system is a technical system or the interaction of different technical systems that have the potential to store energy. Basically, in addition to direct energy storage systems, such as heat or battery storages, inherent storages are possible [21]. Especially in complex production systems, modelling dependencies is particularly important to be able to model flexibility realistically and to limit the flexibility space to the variants that can actually be implemented. The class *flexible load measure* describes a concrete power change of the system within its flexibility space defined by the flexible loads, dependencies, and storages. However, the described model has not yet been used in PPS applications.

Consequently, there is no procedure to potentially plan deployments of energy flexibilities without knowing the actual implementation in advance and at the same time not jeopardizing the achievement of logistical objectives. In addition, there is a lack of a holistic process that meaningfully encompasses the consideration of EF from machine to market and back again. Based on the aforementioned points, the aim of this paper is to present a generic process for the machine-, plant- and market-independent consideration of EF in PPC which considers EF in production planning and provides preconceived options for production control.

3. Methodological approach

With the help of the EFDM, we are developing an approach that allows the integration of EF into existing PPC systems, thus providing an easy-to-use solution for practical applications. Our approach can be divided into five steps as illustrated in Figure 1.

In a first initial step, possible flexibilities are identified and characterized in a flexibility audit before subsequently being modelled in an EFDM. For the identification and characterization of flexibilities, the methodology proposed by Tristan et al. [22] can be used, which examines both production systems directly performing the production tasks and the supply systems performing supporting tasks or tasks necessary for the operation of these production systems. If changes are made to the production infrastructure, processes or product portfolio, a renewed flexibility audit is recommended to leverage the highest possible flexibility potential. After the initial flexibility audit, steps two to five address the operational PPC.

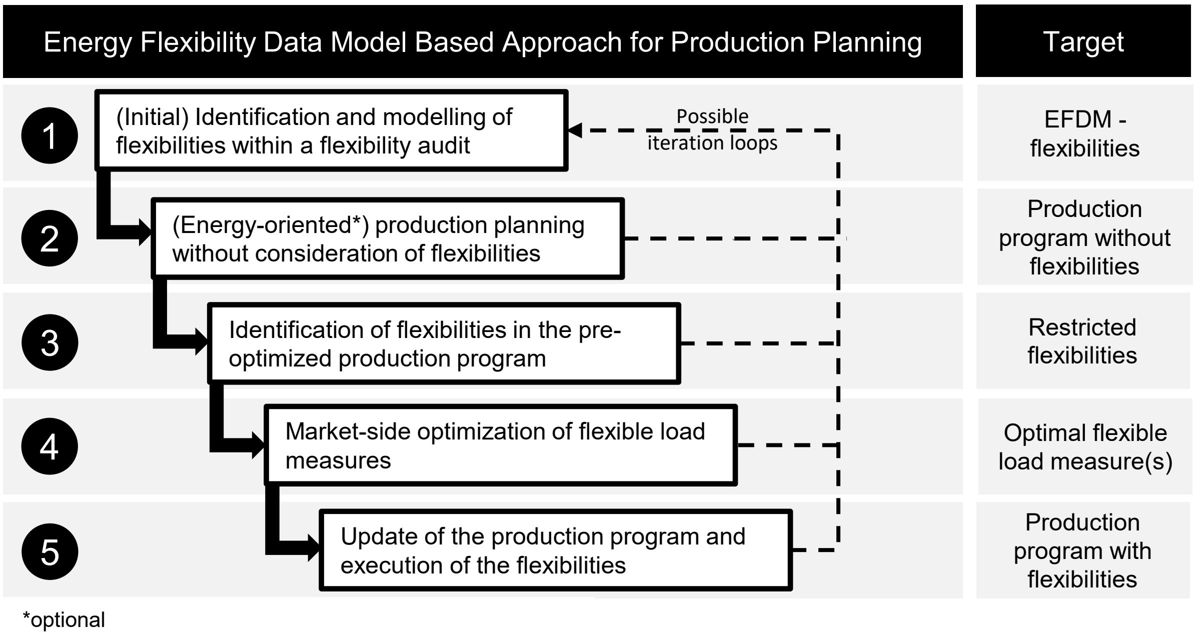


Figure 1: Five-step approach for flexible production planning with an EFDM

In step two, energy-oriented PPC is carried out using conventional software tools without considering the flexibilities previously modelled. An optimisation of logistic objective variables under consideration of a price time series (typically day-ahead prices) for energy-oriented production planning towards a more efficient production is not mandatory for our approach. Conventional PPC software that cannot integrate price time series do not need to be replaced. Our approach is therefore to be understood explicitly as an extension of existing PPC solutions. The result of the second step is consequently a production plan without flexibilities, representing today's standard in advanced production planning systems.

In the third step, the flexibilities modelled in the EFDM are concretized into restricted flexibilities based on the previously created production plan. Restricted flexibilities can only be used and scheduled at positions with free production capacity in the previously created production plan. This procedure ensures that the logistical objectives considered in step two are not violated. In addition, the flexibility space created by the flexibilities modelled in the EFDM is reduced by the restrictions and, as a result, the optimisation carried out in the next step is simplified and more performant.

In the fourth step, the identified restricted flexibilities are optimised regarding the market side. In contrast to energy-oriented production planning, which usually only considers a single price time series of one energy carrier, the nature of flexibilities offers a significantly broader spectrum of marketing options and an associated higher revenue. Thus, in the market-side optimisation the flexible load measures with the highest expected revenues or savings on different possible energy or capacity markets are identified. Here, typically the energy-only markets (day-ahead and intraday) and balancing power markets are integrated in the optimisation [23]. Note here, that the EFDM contains costs of each flexibility (e.g. higher wear of machines or opportunity costs) and allows us to evaluate within the optimisation if the use of flexibilities is advantageous from an economic perspective. Individual costs of flexibilities can be modelled for example with the help of a cost-model initially introduced by Rösch et al. [24] as the quantification of costs for flexibilities in a DSM context might be a challenging task. In addition, the effect of using flexibilities on a company's load profile can also be taken into account, as high load peaks negatively influence the economics by increasing grid charges. Note here, that after the market-side optimisation of flexibilities, marketing steps must be initiated in parallel, which, however, are not the focus of this work. These steps enable the combination of external marketing opportunities and the use of flexibilities for peak shaving. After identifying the optimal flexible load measures to be allocated, the previous production plan from step two can be updated and adjusted with the flexible load measures and associated production tasks determined before the production plan is executed. In the period between the generation of the production plan and its execution, changes in energy and balancing power markets can be monitored and, if necessary, flexible load measures that have already been scheduled can be drawn without major impact on the production plan.

With step five we receive a production plan with flexibilities considered – thus, a more economic production plan. Further iterative loops to dynamically react on machine failures or changes in orders are proposed and applicable for each step allowing for rescheduled production plan.

4. Conceptual application

We demonstrate our five-step approach for energy flexible PPC by conceptually applying it to a use case from a German company in the graphite production. We address both production planning and production control with machine scheduling at the intersection of the two disciplines to optimise the use of EF. Graphitization is the most energy-intensive and last chemical production step in the fabrication of graphite products. The amorphous carbon bodies are heated up to the required temperature of around 2,600 to 3,000°C by resistance heating requiring several megawatts of electrical power [25]. In our case, eight identical furnaces are available for this process step. The furnaces are fed with electrical energy by two transformers, which are moved towards and connected with the specific furnace (see Figure 2 (b)), allowing a maximum of two of them to be operated simultaneously. The heating process is characterized by material specific heating curves and can be divided into two phases (see Figure 2 (a)). In phase one the material is preheated to ensure a constant temperature distribution. After a holding time the target temperature is reached in phase two. Note, one of the transformers has only limited power and therefore cannot cover the entire graphitization process. This transformer can only be used for preheating and thus, its use can partially parallelize two heating processes. For a more detailed process description we refer to Bank et al. [26].

By conceptually applying our approach, in the first step, possible flexibilities have been identified and modelled in an EFDM. The flexibilities identified in the use case can be classified as measures for shifting individual orders and measures for customizing heating curves and thus adapting individual graphitizing processes. For the former, due to the fact that the graphitization process includes a cooling phase, in which the material remains in the furnace, capacities for shifting start times of individual orders are offered in production planning. For the latter, further flexibilities in the form of an adaption of the heating curves are feasible. These include temporary shutdowns as well as load reductions and increases. The flexibilities were modelled in an EFDM and two of them are hereafter used to demonstrate our approach.

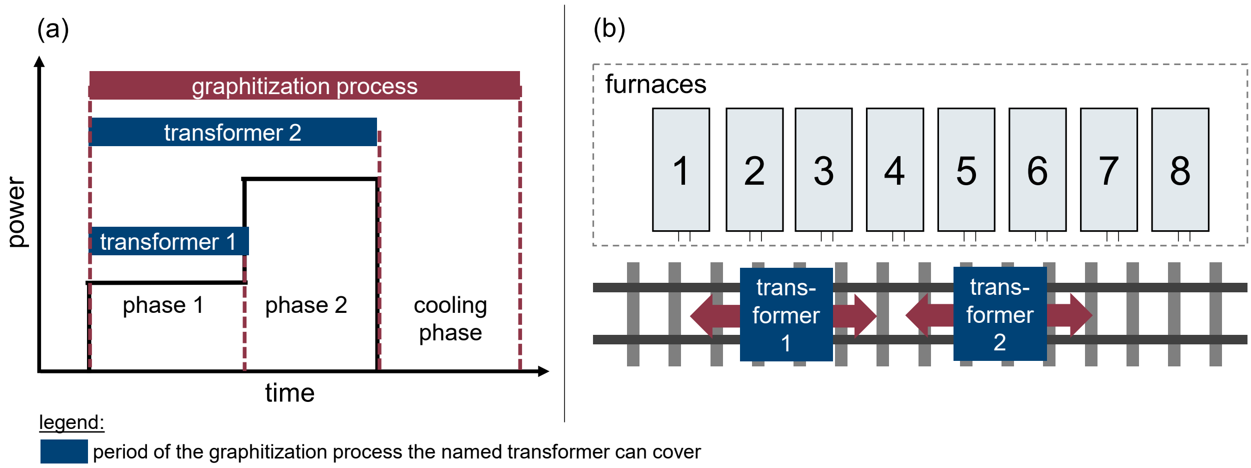


Figure 2: Load curve of the graphitization process (a), layout of the use case (b)

The flexibility “load reduction/shedding by 100% in phase one (flexibility 1 (F1))” can be held for a duration of up to one hour (see Table 1). F1 is coupled with another compensatory flexibility (F2), where (heating) phase two is extended to the extent that compensates for the energy consumption previously missed by F1, thus ensuring that the required total energy demand is met. The flexibilities are not restricted in their validity, i.e. the period in which they can be used. Table 1 shows some additional key figures from the EFDM for the identified flexibilities F1 and F2 exemplarily.

Table 1: Key figures of the identified flexibilities (based on Schott et al. [20])

ID	Description	**F1**	**F2**
validity T_G	period in which the flexibility is available	[00:00, 24:00]	[00:00, 24:00]
reaction duration t_D	time between the command to activate until the start of the change in power	{900s}	{900s}
power states P	power levels during holding periods ((+) increase in power consumption, (-) decrease in power consumption)	{-3 MW}	{+6 MW}
holding duration t_H	length of the period in which the flexible load is held in its power state	{[15min, 60min]}	$-\frac{\int_{i=0}^{T} p_{F1,i} di}{p_{F2,i}}$
dependencies	impact of an activation of a measure on another measure	EFM1 implies EFM2	EFM2 implies EFM1

In step two, energy-oriented production planning is carried out. Therefore, the identified flexibilities are used as degrees of freedom in the optimisation of logistic objective variables under consideration of a day-ahead electricity price prognosis. For this purpose, energy-oriented machine scheduling was implemented in an optimisation tool [26]. Six to seven orders are scheduled for a planning period of seven days in compliance with logistical objectives. The result of the second step is an energy-oriented production plan where orders are planned according to the energy cost forecast without the possibility of additional flexibilities (see Figure 3 (top)).

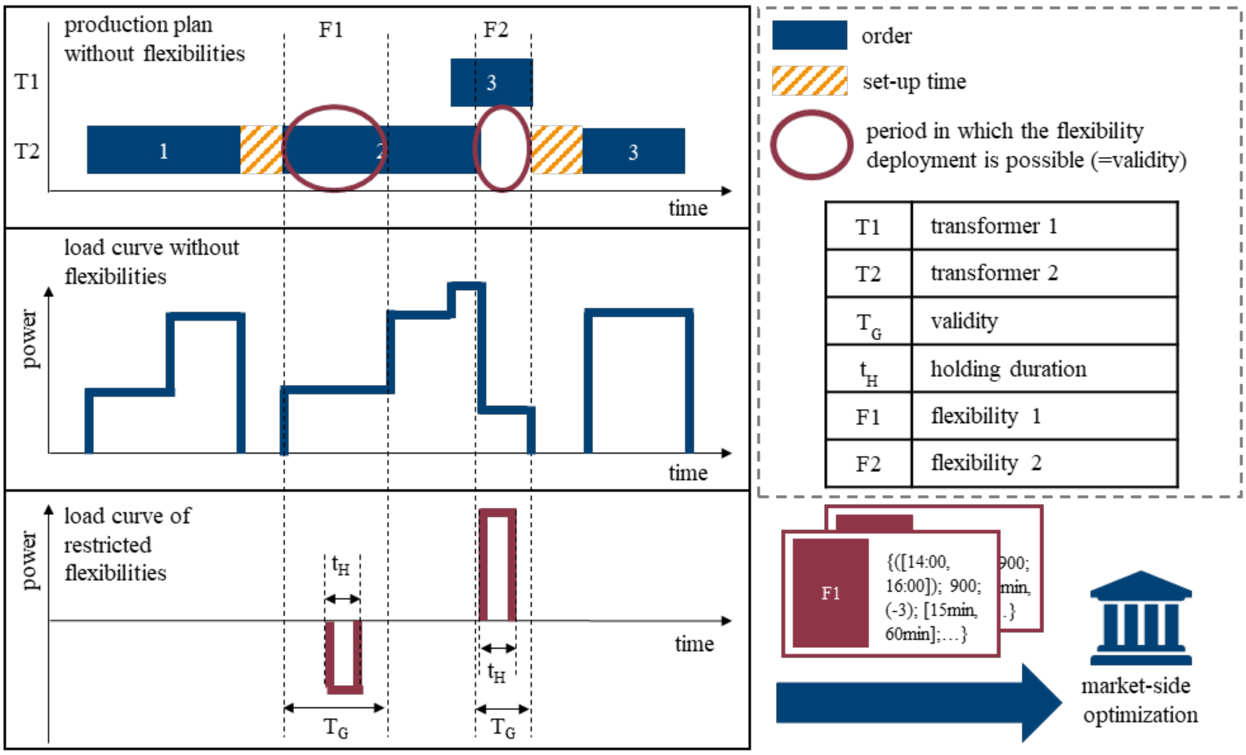


Figure 3: Identification of (restricted) flexibilities in the pre-optimised production plan

In the third step, possible flexibilities in the previously created energy-oriented production plan are identified. For this purpose, the EFDM flexibilities modelled in step one are restricted based on the constraints of the production plan. To ensure that logistical objectives are not affected, flexibilities can only be scheduled in times with free capacity. Key figures of the EFDM such as the holding duration and validity are specified by the restrictions of the production plan. Figure 3 shows the determined production plan (top), its corresponding load curve (middle), and two exemplary flexibilities that can be scheduled based on the available capacities (bottom). The free capacity following order two allows the scheduling of the flexibilities F1 and F2. The holding period of the flexibility spaces is not yet restricted in this step. However, the validity of F1 is restricted to the duration of phase 1 of the heating curve and a specific start time is defined for F2 being the end time of order two.

In the fourth step, the identified restricted flexibilities can be optimised within their remaining degrees of freedom on the market side towards the highest expected revenues of various possible energy or balancing power markets resulting in optimal flexible load measures (FLM) with specific holding durations and starting times. After market side optimisation the production plan from step two can be updated and adjusted with the flexible load measures receiving a production plan with flexibilities considered (see Figure 4). Further iterative loops with dynamically changing planning horizons are applicable. In addition to marketing the (restricted) flexibilities, the pre-planned flexibility options can then also be used in response to the short-term state of production (e.g. in reaction to machine failures) in production control.

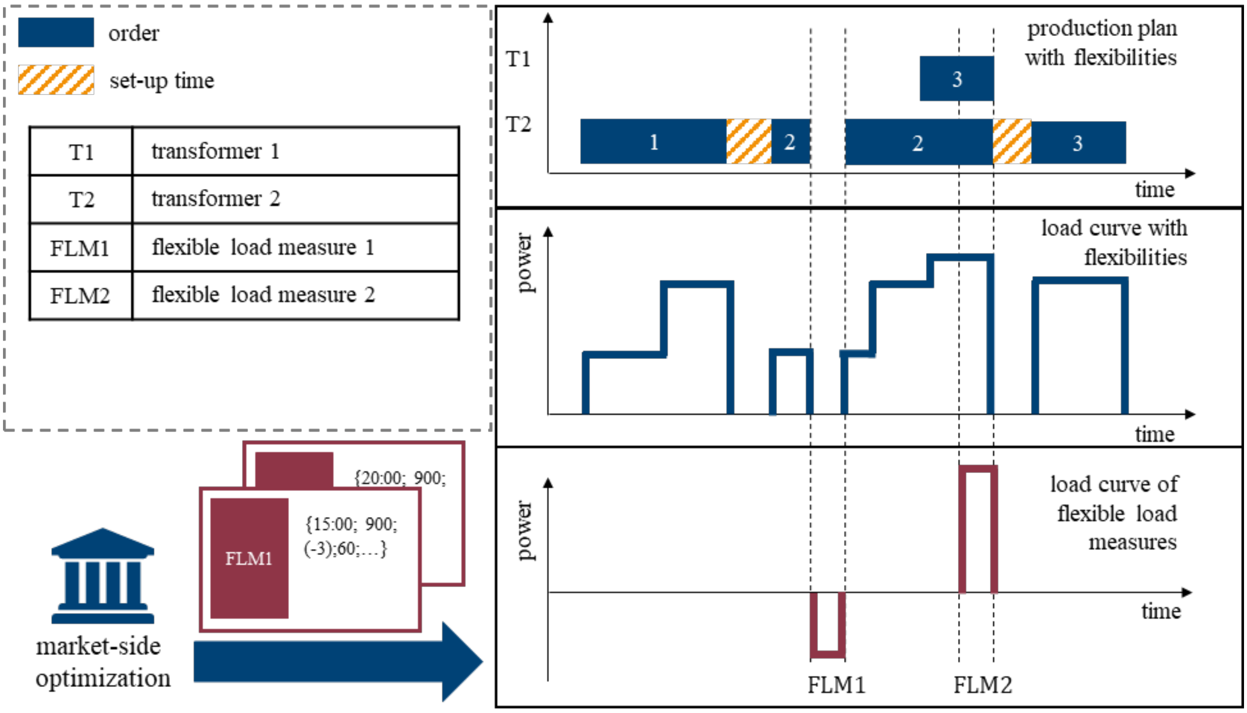


Figure 4: Production plan with flexibilities post market side optimisation

5. Discussion and Outlook

We have developed a five-step approach that considers EF in production planning without compromising the logistical objectives of production. Despite the consideration of EF in planning, only the production control currently decides whether a particular marketing option is taken, based on the corresponding short-term state of production as well as the energy and capacity markets. Thus, the approach closes the existing gap between energy-oriented PPC and flexibility trading as it is described as a building block of a future and sustainable energy system. However, the use case also indicates a limiting factor for practice since flexibility can only be scheduled if free capacities are available. A further monetary advantage can be achieved in addition to the energy-oriented PPC or, in the case of unforeseen events, it is possible to react accordingly and to fall back on flexibility measures that have already been prepared.

The preparation of possible flexibilities closes the gap between production planning and production control which arises in energy-oriented PPC, since different markets and marketing opportunities have different lead times due to individual market designs [27]. Furthermore, due to short-term tradable flexibilities and associated price volatility, deterministic planning of flexibilities in advance is not always possible. In addition, the use of a defined description of EF in the PPC closes the gap between production and energy management because it creates a unified language that can describe flexibility end-to-end from the machine to the energy and capacity markets. To make the procedure more applicable in practice, further automation is necessary. For this purpose, the restricted flexibilities in step three should be automatically generated based on one-time modelled EFDM-flexibilities (step one) and the specific production plan without flexibilities (step two). It is also conceivable to incorporate the revenues to be expected as a result of the flexibility as anticipated values in the planning and to carry out a stochastic optimisation in order to design freedoms in the production plan not only on the basis of the current price curve.

Despite these limitations, with this study we demonstrated a viable approach for integrating EF into existing PPCs, helping to optimise energy costs and therefor help the industry to meet the challenges posed by the energy transition. Our method is particularly suitable due to the step-by-step extension of existing methods and approaches and can also be integrated into approaches that do not yet consider energy as a target variable. In any case, by using the described approach, marketing can take place on different markets, so that a high monetary potential can be achieved in the use of energy flexibility.

Acknowledgements

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] Bundesminesterium für Wirtschaft und Energie, 2019. Kommission „Wachstum, Strukturwandel und Beschäftigung“ Abschlussbericht, Berlin.

[2] Die Bundesregierung. Bundesregierung beschließt Ausstieg aus der Kernkraft bis 2022. https://www.bundesregierung.de/breg-de/suche/bundesregierung-beschliesst-ausstieg-aus-der-kernkraft-bis-2022-457246. Accessed 5 September 2018.

[3] Deutscher Bundestag, 2021. Gesetz für den Ausbau erneuerbarer Energien: EEG.

[4] Graßl, M., 2015. Bewertung der Energieflexibilität in der Produktion. Zugl.: München, Techn. Univ., Diss., 2015. Utz, München, 163 pp.

[5] Verein Deutscher Ingenieure, 2019. VDI-Richtlinie 5207 Energieflexible Fabrik: Grundlagen - Blatt 1. Beuth.

[6] 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.

[7] Keller, F., 2018. Methodik zur energiebezugsorientierten Auftragsplanung. Dissertationsschrift, München.

[8] Akbar, M., Irohara, T., 2018. Scheduling for sustainable manufacturing: A review. Journal of Cleaner Production 205, 866–883.

[9] Wiendahl, H.-P., 1997. Fertigungsregelung: Logistische Beherrschung von Fertigungsabläufen auf Basis des Trichtermodells. Hanser, München, 382 pp.

[10] Rager, M., Gahm, C., Denz, F., 2015. Energy-oriented scheduling based on Evolutionary Algorithms. Computers & Operations Research 54, 218–231.

[11] Heinzl, B., 2020. Methods for Hybrid Modeling and Simulation-Based Optimization in Energy-Aware Production Planning. ARGESIM.

[12] Kaymakci, C., Wenninger, S., Sauer, A., 2021. A Holistic Framework for AI Systems in Industrial Applications. Wirtschaftsinformatik 2021 Proceedings Track 09.

[13] Keller, F., Schultz, C., Braunreuther, S., Reinhart, G., 2016. Enabling Energy-Flexibility of Manufacturing Systems through New Approaches within Production Planning and Control. Procedia CIRP 57, 752–757.

[14] Schultz, C.M., 2018. System zur energieorientierten Produktionssteuerung in der auftragsbezogenen Fertigung. Dissertation, München.

[15] Süße, M., Prell, B., Stoldt, J., Strehle, H.-M., Richter, M., Schlegel, A., Putz, M., 2019. Load and energy management for factories through multi-stage production planning and optimization. Procedia CIRP 80, 257–262.

[16] Wang, H., Jiang, Z., Wang, Y., Zhang, H., Wang, Y., 2018. A two-stage optimization method for energy-saving flexible job-shop scheduling based on energy dynamic characterization. Journal of Cleaner Production 188, 575–588.

[17] Nayak, A., Lee, S., Sutherland, J.W., 2019. Dynamic Load Scheduling for Energy Efficiency in a Job Shop with On-site Wind Mill for Energy Generation. Procedia CIRP 80, 197–202.

[18] Schulz, J., Popp, R.S.-H., Scharmer, V.M., Zaeh, M.F., 2018. An IoT Based Approach for Energy Flexible Control of Production Systems. Procedia CIRP 69, 650–655.

[19] 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.

[20] Schott, P., Sedlmeir, J., Strobel, N., Weber, T., Fridgen, G., Abele, E., 2019. A Generic Data Model for Describing Flexibility in Power Markets. Energies 12 (10), 1893.

[21] Strobel, N., Fuhrländer-Völker, D., Weigold, M., Abele, E., 2020. Quantifying the Demand Response Potential of Inherent Energy Storages in Production Systems. Energies 13 (16), 4161.

[22] Tristán, A., Heuberger, F., Sauer, A., 2020. A Methodology to Systematically Identify and Characterize Energy Flexibility Measures in Industrial Systems. Energies 13 (22), 5887.

[23] Wang, H., Riaz, S., Mancarella, P., 2020. Integrated techno-economic modeling, flexibility analysis, and business case assessment of an urban virtual power plant with multi-market co-optimization. Applied Energy 259, 114142.

[24] Roesch, M., Berger, C., Braunreuther, S., Reinhart, G., 2018 - 2018. Cost-model for Energy-oriented Production Control, in: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok. 16.12.2018 - 19.12.2018. IEEE, pp. 158–162.

[25] Jäger, H., Frohs, W., Banek, M., Christ, M., Daimer, J., Fendt, F., Friedrich, C., Gojny, F., Hiltmann, F., Meyer zu Reckendorf, R., Montminy, J., Ostermann, H., Müller, N., Wimmer, K., Sturm, F. von, Wege, E., Roussel, K., Handl, W., 2000. Carbon, 4. Industrial Carbons, in: , Ullmann's Encyclopedia of Industrial Chemistry, vol. 55. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, p. 1.

[26] Bank, L., Rösch, M., Unterberger, E., Roth, S., Rohrer, A., Köberlein, J., Braunreuther, S., Schilp, J., 2019. Comparison of Simulation-based and Optimization-based Energy Flexible Production Planning. Procedia CIRP 81, 294–299.

[27] Seitz, P., Abele, E., Bank, L., Bauernhansl, T., Colangelo, E., Fridgen, G., Schilp, J., Schott, P., Sedlmeir, J., Strobel, N., Weber, T., 2019. IT-based Architecture for Power Market Oriented Optimization at Multiple Levels in Production Processes. Procedia CIRP 81, 618–623.

Biography

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 optimisation and simulation are also related to this.

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 and focuses on research in the fields of data analytics in an industrial and energy context.

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

Martin Lindner is working at the Technical University of Darmstadt, Institute of Production Management, Technology and Machine Tools (PTW) since May 2019. His research is focused on the simulation, optimization and flexibilization of energy systems within the production infrastructure using modern data communication standards and machine learning methods.

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.

Prof. Dr.-Ing. Matthias Weigold is the Director of the Institute for Production Management, Technology and Machine Tools (PTW) at the Technical University of Darmstadt and focus his research on digitalisation and sustainability in production.

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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Customer Success Management In The Subscription Business Of Manufacturing Companies: Towards A Task-Oriented Reference Model

Lennard Holst[1], Günther Schuh[1], Volker Stich[1], Jana Frank[1]
[1]Institute for Industrial Management (FIR) at RWTH Aachen University, Aachen, Germany

Abstract

Manufacturing companies are constantly increasing their efforts in the subscription business, also known as product-as-a-service business, offering usage and outcome based solutions (value-in-use) instead of transactional services and products (value-in-exchange). Customers are becoming contractual subscribers of the solution in return for recurring, performance-related payments. To address arising, inevitable challenges like (1) reducing customer churn, (2) increasing usage intensity and outcome quality, (3) ensuring the adoption of product and software releases as well as (4) fostering customer loyalty, leading manufacturing companies are setting up a new organizational, customer-facing unit, called Customer Success Management (CSM). This unit has its origins in the software-as-a-service business, operating next to established entities like sales, key account management and customer service. Since there are currently no holistic models for an end-to-end description of CSM-tasks in the manufacturing industry, this paper contributes to a task-oriented reference model, using a grounded theory approach, examining both manufacturing and software companies. Containing a reference framework with 8 main tasks, 17 basic tasks and 76 elementary tasks, the reference model supports manufacturing companies in adapting and customizing a company-specific CSM-concept.

Keywords

Customer Success Management; Subscription Business Models; Value-in-use, Manufacturing Companies; Task-oriented Reference Model

1. Introduction

Economic growth and competitive differentiation in the manufacturing sector are no longer based on product innovations and digital services alone but on the ability to monetize the usage phase of products in customer operations [1]. Due to market saturation and interchangeability of products, manufacturing companies are increasingly examining entry and growth opportunities in the subscription business as a way of supplementing the conventional sale of products and services [2]. In the subscription business, instead of a product or service, the productive state in the customer's operation is paid for [3]. In the following, we subsume related terms such as product-as-a-service, everything-as-a-service or pay-per-x-models under the term subscription business model, as this term puts the focus on the underlying shift from a product-centric to a fully customer- or subscriber-centric business approach [4,5,1]. While subscription business models, especially in the software industry, are often narrowly defined by the attribute of periodic payments in return for a recurring service, the possibilities of digitization lead to a broader definition of the term subscription in the industrial context [6,7]. Thus, the definition of the term subscription in this paper is based on four key characteristics [8], as shown in Table 1.

Table 1: Subscription characteristics and derived requirements for a Customer Success Management

Subscription business model characteristics		Requirements for Customer Success Management
-1- Recurring revenue mechanisms such as pay-per-use or pay-per-outcome	→	Reducing financial risks by preventing customer churn and increasing customer retention
-2- A value proposition geared to individual customer success (e.g. increasing productivity)	→	Monitoring customer operations, increasing usage intensity and outcome quality
-3- Digital performance systems networked via the Internet of Things, consisting of integrated bundles of products, services and software	→	Ensuring continuous onboarding, training, performance consulting and the adoption of product and software releases
-4- Long-term customer relationship based on a trustworthy partnership	→	Fostering customer loyalty and a positive lock-in-effect while becoming the "trusted advisor"

From a research perspective, the subscription business can be interpreted as the operationalization of the value-in-use concept proposed by VARGO A. LUSCH: Instead of delivering value-in-exchange by selling products or services on a transactional basis, suppliers align their economic success to the continuous value realization or co-value-creation in the usage phase of products and services [9,10].

However, despite the overarching opportunities of a deep customer relationship, data-based understanding of changing customer needs and economic growth after overcoming a period of starting losses due to smaller recurring revenues over time, subscription businesses are associated with numerous challenges, especially through the proportional assumption of risks from the customer processes [11]. This places new demands on customer management in particular [12], as shown in Table 1. Existing customer-facing business concepts or units, such as customer relationship management, customer service, sales or key account management are often holistically not designed and incentivized for these new proactive, analytical tasks to ensure and increase customer success [13] Therefore the subscription business requires the setup of a new customer-facing business approach or unit called Customer Success Management (CSM) [14].

In contrast to the more subjective terms of value, value realization and co-value-creation, we define *customer success (CS) as the measurable achievement of economic, factual, environmental or person-related individual customer objectives*. Based on this definition, Customer Success Management, whether as a business concept or as a separate unit, monitors and analyzes the usage phase of the products and services and proactively supports the subscription customers in achieving and increasing their individual objectives.

While Customer Success Management has already been established in the software industry for several years, companies in the manufacturing sector are still in the conceptual phase of Customer Success Management, parallel to the setup and expansion of their subscription business [14]. Since no descriptions of CSM reference tasks for manufacturing companies are available so far, this paper aims to answer the following research question: How does a reference model need be designed to describe the end-to-end tasks of a Customer Success Management for manufacturing companies? In the following chapter, the need for research is identified based on a literature review, before the specific approach to answer the research question is derived. As result of the research process, we present the task-oriented CSM-reference model. Finally, the limitations and the resulting need for further research are discussed.

2. State of research

In recent scientific literature, first contributions with the CSM as the subject of investigation exist, mostly against the background of the software industry and thus without direct reference to specific tasks of a CSM in the manufacturing sector. EGGERT ET AL. provide a differentiation of CSM from other management concepts such as quality, customer satisfaction, key account, and customer experience management by emphasizing CSM competencies to proactively engage in the customer experience processes [15]. The authors focus on the development of a scientific perspective on CSM and the conceptual adaptation of CSM from the IT context to other industries. However, there is no detailed description of the implications and CSM specific tasks for the manufacturing industry. HILTON ET AL. name the three necessary research fields Goal Management, Learning Management and Stakeholder Management for investigating Customer Success Management [16]. They refer mostly to the software industry by naming practical examples. There is no explicit consideration of the implications for manufacturing companies. HOCHSTEIN ET AL. look at CSM developments from a combined service and sales perspective [17]. The authors describe CSM as a way of breaking down the service-sales ambidexterity, i.e. selling through service and carrying out service activities through sales, so as not to overburden the customer with too many supplier contacts. In this context, the authors see CSM as a fusion of existing concepts such as customer centricity, management of customer touchpoints and after-sales service activities in a new function. While discussing specific CSM tasks that are relevant for manufacturing companies, the contribution does not provide a referenceable model for CSM tasks. PORTER A. HEPPELMANN examine the changes in business activities and organizational structures in companies due to the development of intelligent, networked products [14]. The authors name selected tasks of the CSM, such as the responsibility for the customer experience as well as the data-related monitoring and assurance of value realization by the customer. The article can be directly applied to manufacturing companies, although it does not yet provide a holistic, task-oriented CSM reference model. ADAMS provides a comprehensive CSM model with his management-oriented contribution, which, however, does not directly refer to the requirements of the manufacturing industry [18].

With regard to the scientific literature in the area of subscription business models, the requirements for a CSM are named, but neither CSM tasks are described nor referenceable processes are designed. SCHUH ET AL. refer to new organizational functions and tasks within the subscription business and explicitly name Customer Success Management as the unit responsible for the ongoing customer relationship and for generating value in the customer process [2]. Furthermore, the authors examine the digital shadow within the IT-reference framework Internet of Production as the data basis for carrying out CSM tasks. STOJKOVSKI ET AL. primarily compare the advantages and risks of the subscription business for suppliers and customers in the manufacturing sector, without explicitly addressing the CSM [1]. TZUO A. WEISERT as well as LAH provide current standard works describing the transformation from product-centric to customer-centric companies in the course of the so-called subscription economy [12,3]. Customer Success Management plays a central role in their contributions, since success in the subscription business will continue as long as customers use the services successfully. To this end, the authors demand that the daily use of the service, customer satisfaction as well as specific downtimes must be monitored. The article is primarily written with reference to the software industry. However, the authors repeatedly make reference to the manufacturing industry, so that the models developed for CSM can be partially applied in manufacturing contexts.

Regarding the scientific literature in the field of reference modelling, to this date no explicit models for CSM tasks can be found. For example, BECKER ET AL. contrast numerous existing reference models, none of which explicitly encompasses the tasks of a CSM, neither in software nor in manufacturing contexts [19]. For this reason, we consider the current research gap of a task-oriented reference model for Customer Success Management in manufacturing companies as persisting.

3. Methodology

Since the Customer Success Management is a new management approach or business unit in the manufacturing industry and since the results of this research are intended to be directly suitable for reuse in the form of a reference model, an applied research approach is chosen for data collection and analysis. Therefore, we use the Grounded Theory according to GLASER A. STRAUSS to systematically collect data as a theoretical sample, identify recurring elements and derive a theoretical reference model that is finally tested and validated in practice [20]. In this approach, different research methods for collecting and analyzing primary data are combined, such as expert interviews, observations and focus groups. In addition, the results of the literature analysis will be used to enrich the reference model with secondary data. Table 2 provides an overview of the applied research process that took place between the years 2019 and 2021.

Table 2: Overview of the applied research methods with company and interview partners

Company	CSM status	Interview partner role	Applied research methods
Machine tool manufacturer	CSM in preparation	Head of Subscription Head of Process Management	3 interviews, 1 company visit, 2 joint workshops between 2019 and 2020
Printing press manufacturer	CSM in action	Head of Customer Success Management Operative CS Manager	4 interviews, 1 observation of a customer success meeting with customers between 2019 and 2021
Machine tool manufacturer	CSM in action	Operative CS Manager	1 interview in 2020
Software as a Service	CSM in action	Director Marketing Campaigns & Channels	3 interviews, 1 joint workshop in 2020
Software as a Service	CSM in action	Success Management Lead	2 interviews in 2019
Various manufacturing companies	CSM in preparation	Various roles	Focus group in the context of a joint remote workshop in 2020

4. Results

In this section, the task-oriented CSM reference model is developed. At first, a regulatory framework of the reference model is presented in Figure 1, serving as a top-level guidance tool through the reference model.

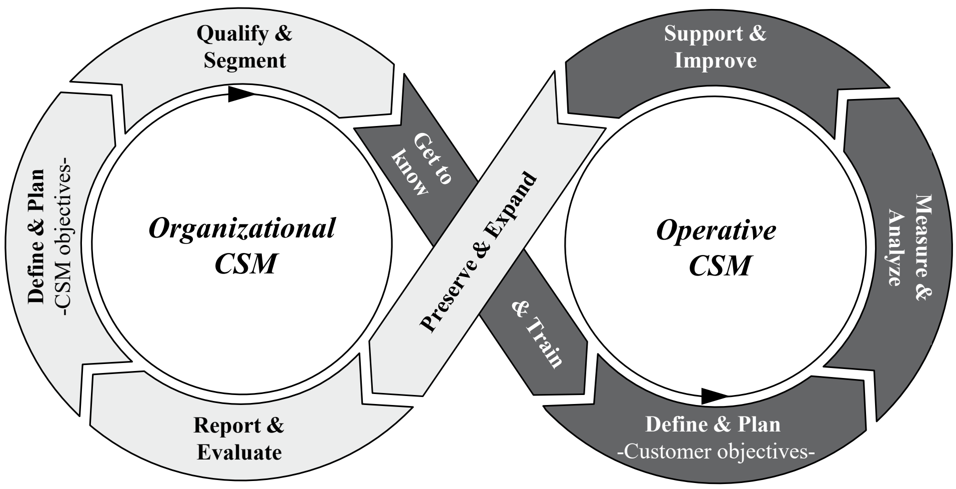


Figure 1: Regulatory framework of the task-oriented CSM reference model

On the one hand, the framework presented in Figure 1 consists of eight successive, order-giving dimensions, which at the same time represent the main tasks of the CSM reference model. Since the core function of CSM is the continuous improvement of customer processes to ensure individual customer success, the DMAIC-cycle [21] was used in combination with the concept of the customer journey [22] to derive the eight order-giving dimensions of CSM tasks. On the other hand, the graphical representation as a "lying eight" illustrates the cyclical character of the CSM tasks and divides the CSM reference model into two levels, in analogy to the St. Galler Key Account Management model [23]. The organizational CSM subsumes preparatory tasks of the top-level management for the creation of the organizational prerequisites for a successful CSM. The operational CSM subsumes the daily tasks of a customer success manager or a business unit entrusted with implementing CSM activities. In the following, further basic and elementary tasks are assigned to the eight main tasks shown in the regulatory framework.

4.1 Define & Plan (CSM objectives)

CSM must be positioned and aligned organizationally in the company of the CSM provider. To this end, the internal objectives of CSM and the human and technological resources required to achieve them must be defined.

Table 3: Basic and elementary tasks assigned to the first main task "Define & Plan"

Main task	Basic task	Elementary tasks	
1 Define & Plan *(CSM objectives)*	1.1 CSM objective definition	1.1.1 Motivate CSM invest 1.1.2 Define CSM terms 1.1.3 Establish CSM strategy and CSM culture	1.1.4 Define CSM objectives 1.1.5 Define CSM KPIs 1.1.6 Define KPI threshold values
	1.2 Organization and resource planning	1.2.1 Define CSM tasks 1.2.2 Set up organizational structure 1.2.3 Choose and incentivize CSM employees	1.2.4 Choose and integrate IT-systems 1.2.5 Ensure financial funding of CSM activities

4.2 Qualify & Segment

The CSM must ensure that subscription relationships are entered into with potential customers for whom there is an opportunity to guarantee or increase customer success. In addition to qualifying customers, they should be divided into individual CSM-specific customer segments, such as low touch or high touch customer segments with varying levels of personal interaction efforts [24].

Table 4: Basic and elementary tasks assigned to the first main task "Qualify & Segment"

Main task	Basic task	Elementary tasks	
2 Qualify & Segment	2.1 Customer qualification	2.1.1 Define criteria for potential subscription customers 2.1.2 Analyze customer production performance 2.1.3 Analyze customer business model	2.1.4 Perform risk analysis of submitting contract offer 2.1.5 Manage expectations towards the subscription services
	2.2 Customer segmentation	2.2.1 Define criteria for customer segments (e.g. low touch or high touch customer segments) 2.2.2 Segment the customers according to the set-up criteria	2.2.3 Designate & prepare CSM responsible entities 2.2.4 Define vendor's subscription team (e.g. CSM as the subscription project lead)

4.3 Get to Know & Train

Parallel to the set up of the subscription services in the customer process, usually carried out by the technical service team, the CSM must establish contact with the customer's usage center and introduce and explain central, digital service components and functions [24]. The CSM must ensure that the customer derives measurable benefits from the subscription service as quickly as possible. This main task represents the transition from organizational to operational Customer Success Management.

Table 5: Basic and elementary tasks assigned to the first main task “Define & Plan”

Main task	**Basic task**	**Elementary tasks**	
3 Get to Know & Train	3.1 Stakeholder-management	3.1.1 Establish customer contact 3.1.2 Conduct target workshop with customer	3.1.3 Define customer’s subscription team 3.1.4 Analyze customer`s subscription team
	3.2. Installation support and training management	3.2.1 Plan installation phase of subscription services 3.2.2 Explain and configure the individual subscription performance system 3.2.3 Develop training plans	3.2.4 Conduct training courses 3.2.5 Automate the onboarding process 3.2.6 Monitor the installation & training sessions

4.4 Define & Plan (Customer objectives)

Together with the customer, the CSM determines the individual customer success criteria and plans the achievement of specific customer objectives, such as the systematic increase in customer productivity over a defined period. Furthermore, the CSM plans the specific measures to achieve the goals together with the customer and sets up an individual customer success roadmap.

Table 6: Basic and elementary tasks assigned to the first main task “Define & Plan”

Main task	**Basic task**	**Elementary tasks**	
4 Define & Plan *(Customer objectives)*	4.1 Customer objective definition	4.1.1 Define customer objectives 4.1.2 Define customer KPIs 4.1.3 Define KPI threshold values	4.1.4 Manage expectations towards the CSM services
	4.2. Customer objective planning	4.2.1 Develop customer success roadmap 4.2.2 Prepare performance consulting 4.2.3 Establish process evidence	4.2.4 Perform risk analysis of customer objective achievement

4.5 Measure & Analyze

The CSM must measure and analyze the available data from the utilization phase in order to be able to intervene in the customer process in the event of critical deviations or to improve the process proactively. The core concept within this dimension is setting up a Customer Health Score (CHS) measurement. The CHS consists primarily of leading KPIs related to product and service usage in customer operations. Metrics that can be tracked by the CSM are, e.g., the active time in applications, the realization of customer success criteria, customer satisfaction scores, or open service tickets. Continuous monitoring of the CHS allows a proactive response to changes or problems in customer operations.

Table 7: Basic and elementary tasks assigned to the first main task “Define & Plan”

Main task	Basic task	Elementary tasks	
5 Measure & Analyze	5.1 Customer Health Scoring (CHS)	5.1.1 Motivate and define CHS 5.1.2 Determine metrics for CHS 5.1.3 Develop CHS evaluation system	5.1.4 Determine threshold values for the CHS 5.1.5 Integrate & visualize CHS
	5.2 Customer objective controlling	5.2.1 Record & analyze CHS 5.2.2 Capture & analyze customer KPIs 5.2.3 Compare CHS & KPIs with benchmark	5.2.4 Perform root cause analysis in case of deviations 5.2.5 Derive action items & initiate internal improvements
	5.3 Feedback management	5.3.1 Stimulate customer feedback 5.3.2 Accept & analyze customer feedback	5.3.3 Perform root cause analysis in case of deviations 5.3.4 Derive & report measures

4.6 Support & Improve

Through proactive performance consulting and reactive first-level support, the CSM must ensure the realization of individual customer success so that the subscription provider can generate long-term revenue with the customer or grow together with the customer.

Table 8: Basic and elementary tasks assigned to the first main task “Define & Plan”

Main task	Basic task	Elementary tasks	
6 Support & Improve	6.1 First Level Support	6.1.1 Record problem 6.1.2 Identify & classify cause of problem 6.1.3 Solve problem immediately	6.1.4 Forward problem to 2nd or 3rd level support 6.1.5 Track problem resolution 6.1.6 Document resolution
	6.2 Performance consulting	6.2.1 Prepare individual performance consulting meeting 6.2.2 Present customer KPIs 6.2.3 Present action plan	6.2.4 Celebrate realized customer success 6.2.5 Present new releases 6.2.6 Document & share results

4.7 Preserve & Expand

The CSM is responsible for preserving and expanding the customer relationship. To this end, it identifies potential for expansion of the customer account that will generate mutual benefits. Thus, the CSM increases the customer lifetime value in the best interest of both customer and vendor [25]. This main function represents the transition from operational to organizational Customer Success Management.

Table 9: Basic and elementary tasks assigned to the first main task “Define & Plan”

Main task	Basic task	Elementary tasks	
7 Preserve & Expand	7.1 Churn Management	7.1.1 Determine factors influencing customer churn behaviour 7.1.2 Identify unstable subscribers	7.1.3 Initiate & monitor stabilization measures
	7.2 Contract & expansion management	7.2.1 Monitor & adjust contract status of subscribers 7.2.2 Proactively increase the customer lifetime value	7.2.3 Accompany contract renewals & terminations

4.8 Report & Evaluate

The CSM should systematically share the gained knowledge about the individual customers within the own organization and evaluate the CSM efforts and the fulfillment of CSM KPIs [26]. This task completes the CSM loop, presented in the regulatory framework.

Table 10: Basic and elementary tasks assigned to the first main task "Define & Plan"

Main task	Basic task	Elementary tasks	
8 Report & Evaluate	8.1 Knowledge management	8.1.1 Formalize knowledge about subscribers 8.1.2 Expand knowledge about subscribers from further channels	8.1.3 Proactively share knowledge about subscribers (e.g. via articles in the intranet or knowledge sharing events)
	8.2 CSM objective controlling	8.2.1 Analyze CSM KPIs 8.2.2 Identify causes for derivations and derive measures	8.2.3 Ensure cost-effectiveness of CSM activities

4.9 Application of the CSM reference model in practice

The essential feature of a reference model is the reuse for the derivation of company-specific models [27]. For this reason, the CSM reference model was tested in three selected use cases with companies from the tool, kitchen, and plant engineering sector. In all three use cases, the reference model with its recommendations was applied by a company representative from a subscription- or CSM-related business department to derive and configure a company-specific CSM model. Thus, in the sense of critical rationalism, the hypothesis of reusability is upheld and the model is considered valid for practical reuse and for further adjustments in the course of future developments in CSM research.

5. Conclusion

In this paper, a research process was described for deriving a CSM reference model for manufacturing companies that enter or expand subscription business activities, thus aligning with the realized success of their customers. Considering the identified research gap of a holistic task-oriented CSM reference model, a framework with 8 main tasks of the CSM in manufacturing contexts was derived on the basis of an applied research approach. The 8 main tasks were further detailed by describing 17 basic tasks and 76 elementary tasks. The CSM reference model was finally tested in practical use cases and was found to be reusable by the test users. By its very nature, the CSM reference model in this paper only represents an excerpt in the temporal course of developments regarding the subscription business and the CSM. Thus, there is a need for further development and research regarding process and data modelling, continuous quality assurance, and further adaptation of the reference model. Additionally, there is a need for further research regarding the measurement of quantitative effects of the CSM in order to capture and prove the long-term benefits of the CSM.

Acknowledgements

The research and development project *Future Data Assets* that forms the basis for this report is funded within the scope of the 'Smart Data Economy' technology program run by the German Federal Ministry for Economic Affairs and Energy (BMWi) and is man-aged by the DLR project management agency. The authors are responsible for the content of this publication.

References

[1] Stojkovski, I., Achleitner, A.-K., Lange, T., 2021. Equipment as a service: The transition towards usage-based business models. https://ssrn.com/abstract=.

[2] Schuh, G., Frank, J., Jussen, P., Rix, C., Harland, T., 2019. Monetizing Industry 4.0: Design Principles for Subscription Business in the Manufacturing Industry, in: Proceedings, 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). Sophia Antipolis Innovation Park, France, 17-19 June 2019. 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), Valbonne Sophia-Antipolis, France. IEEE, Piscataway, NJ, pp. 1–9.

[3] Tzuo, T., Weisert, G., 2018. Subscribed: Why the subscription model will be your company's future - and what to do about it.

[4] Claasen, M., Blum, C., Osterrieder, P., Friedli, T., 2019. Everything as a service?: Introducing the St.Gallen IGaaS Management Model, in: Meierhofer, J., West, S. (Eds.), 2nd Smart Services Summit, Thun, pp. 61–65.

[5] Mansard, M., Cagin, J.-M., 2020. Reaping The Recurring Benefits of Industry 4.0: A manufacturing executive playbook for business model transformation towards new revenue streams. Zuora, Roland Berger.

[6] Gassmann, O., Frankenberger, K., Csik, M., 2014. The business model navigator: 55 models that will revolutionise your business. Pearson, Harlow, England, London, New York, Boston, San Francisco, 387 pp.

[7] Mell, P., Grance, T., 2011. The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology. Special Publication 800-145. National Institute of Standards and Technology, Gaithersburg.

[8] Schuh, G., Wenger, L., Stich, V., Hicking, J., Gailus, J., 2020. Outcome Economy: Subscription Business Models in Machinery and Plant Engineering. Procedia CIRP 93 (1), 599–604.

[9] Vargo, S.L., Lusch, R.F., 2004. Evolving to a New Dominant Logic for Marketing. Journal of Marketing 68 (1), 1–17.

[10] Vargo, S.L., Lusch, R.F., 2008. From goods to service(s): Divergences and convergences of logics. Industrial Marketing Management 37 (3), 254–259.

[11] Stoppel, E., Roth, S., 2017. The conceptualization of pricing schemes: From product-centric to customer-centric value approaches. J Revenue Pricing Manag 16 (1), 76–90.

[12] Lah, T., Wood, J.B., 2016. Technology-as-a-Service Playbook: How to Grow a Profitable Subscription Business. Point B Inc, [United States].

[13] Mehta, N., Steinman, D., Murphy, L., 2016. Customer Success: How Innovative Companies Are Reducing Churn and Growing Recurring Revenue, 1. Aufl. ed. Wiley, s.l., 256 pp.

[14] Porter, M.E., Heppelmann, J.E., 2015. How Smart, Connected Products Are Transforming Companies. Harvard Business Review 93, 1–19. https://hbr.org/2015/10/how-smart-connected-products-are-transforming-companies.

[15] Eggert, A., Ulaga, W., Gehring, A., 2020. Managing Customer Success in Business Markets: Conceptual Foundation and Practical Application. SMR 4 (2-3), 121–132.

[16] Hilton, B., Hajihashemi, B., Henderson, C.M., Palmatier, R.W., 2020. Customer Success Management: The next evolution in customer management practice? Industrial Marketing Management 90, 360–369.

[17] Hochstein, B., Rangarajan, D., Mehta, N., Kocher, D., 2020. An Industry/Academic Perspective on Customer Success Management. Journal of Service Research 23 (1), 3–7.

[18] Adams, R., 2019. Practical customer success management: A best practice framework for rapid generation of customer success, First edition ed. Routledge, New York, NY, 245 pp.

[19] Becker, J., Beverungen, D.F., Knackstedt, R., 2010. The challenge of conceptual modeling for product–service systems: status-quo and perspectives for reference models and modeling languages. Inf Syst E-Bus Manage 8 (1), 33–66.

[20] Glaser, B.G., Strauss, A.L., 1967. The discovery of grounded theory: Strategies for qualitative research, 4. paperback printing ed. Aldine, New Brunswick, 271 pp.

[21] Larson, A., 2003. Demystifying six sigma: A company-wide approach to continuous improvement. AMACOM, New York, 191 pp.

[22] Lemon, K.N., Verhoef, P.C., 2016. Understanding Customer Experience Throughout the Customer Journey. Journal of Marketing 80 (6), 69–96.

[23] Belz, C., Müllner, M., Zupancic, D., 2010. Excellence in Key Account Management: The St. Gallen KAM concept, 2., updated and rev. ed. ed. mi-Wirtschaftsbuch FinanzBuch-Verl., München, 348 pp.

[24] Naisbitt, J., Naisbitt, N., Philips, D., 2001. High tech high touch: Technology and our accelerated search for meaning, Paperback ed., reprinted. ed. Brealey, London, 274 pp.

[25] Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., Sriram, S., 2006. Modeling Customer Lifetime Value. Journal of Service Research 9 (2), 139–155.

[26] Nonaka, I., Toyama, R., Konno, N., 2000. SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation. Long Range Planning 33 (1), 5–34.

[27] Rehse, J.-R., Fettke, P., 2019. Clustering Business Process Activities for Identifying Reference Model Components, in: Daniel, F., Sheng, Q.Z., Motahari, H. (Eds.), Business Process Management Workshops, vol. 342. Springer International Publishing, Cham, pp. 5–17.

Biography

Lennard Holst, M. Sc., is group leader of the research unit Subscription Business Management within the department of Service Management at the Institute for Industrial Management (FIR) at RWTH Aachen University. He is in charge of various applied research projects regarding the digital transformation of manufacturing companies from product-centric to customer-centric business models.

Prof. Dr.-Ing. Dipl.-Wirt. Ing. Günther Schuh is director of the Institute for Industrial Management (FIR) at the RWTH Aachen, holds the Chair of Production Systems at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University and is a member of the board of directors of the Fraunhofer Institute for Production Technology IPT. He is the founder of Schuh & Co. Group and the e.GO Mobile AG. Prof. Schuh is a member of several supervisory boards and boards of directors.

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

Dr.-Ing. Jana Frank has been head of the department of Service Management at the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2019. Dr.-Ing. Jana Frank leads various research and industrial projects in the fields of subscription business management, digital business model transformation as well as industrial service management.

2nd Conference on Production Systems and Logistics

Linked Accomplishment Of Order Management And Production Planning And Control. An Integrated Model-based Approach

Simon Hillnhagen[1], Alexander Mütze[2], Peter Nyhuis[2], Matthias Schmidt[1]

[1]*Institute of Product and Process Innovation (PPI), Leuphana University of Lüneburg, Lüneburg, Germany*
[2]*Institute of Production Systems and Logistics (IFA), Leibniz University Hannover, Hanover, Germany*

Abstract

Order management is a main PPC task in company internal supply chains. It provides the link between customers and the company's internal order processing processes. The boundary conditions defined hereby influence adjacent and subsequent main tasks in PPC. However, since usually conflicts between objectives, such as between utilisation of capacities and schedule reliability are caused in companies, the interaction between PPC tasks of order management and other main PPC tasks should be taken into account.

Existing approaches often examine the effect of PPC tasks and procedures in isolation for the respective main task. Especially in the task of order management, the focus of procedures and models lies more on economic and less on production logistics aspects. Nevertheless, these must be taken into account to assess the resulting effects on production (e.g. load and due-date situation) and thus on the logistics performance towards the customer.

To counteract this, the contribution presents a model-based approach, which considers the interactions between order management tasks and other PPC tasks. Furthermore, the determined interdependencies and links are used to derive the effects of order management on logistics objectives of production, and to identify how an integrated view can be used to raise production logistics potentials for improving a company's logistics performance. For this purpose, proven models such as the Hanoverian Supply Chain Model and the extended Aachen PPC Model are used.

Keywords

Production Planning and Control; PPC; Order Management; Configuration; Model

1. Introduction

Companies face the challenge of positioning themselves in the trade-off between high logistics performance and low logistics costs in order to keep their competitiveness and ultimately their profitability high [1]. To achieve this, it is essential to configure the processes along the company's internal supply chain, and in particular production planning and control (PPC), as a key lever, with its tasks and procedures in a targeted and comprehensive manner [2,3]. Despite the high importance of PPC, which is also expressed in a large number of research publications, studies show that the understanding of processes in PPC, as well as the degree of utilisation of the available systems, is still needing improvement. For example, two studies in Germany conducted by research institutes, in which specialists and managers were surveyed, attest to the fact that only 20% of employees rate their understanding of the PPC system as high [4], and three-quarters of the respondents attest their employees no to moderate knowledge regarding the process effect in the context of production control [5]. Furthermore, only about 10% of the respondents state that the potential of

PPC systems is fully utilised in their company [6]. Considering these study results, it is not surprising that about two-thirds of the surveyed companies from the German mechanical and plant engineering sector are not satisfied with the organisation of their production planning and see high potential for improvement in the achievement of production logistics objectives, such as schedule reliability [7].

Based on the above-mentioned studies, it can be concluded that, on the one hand, there is still potential for improvement in PPC, but on the other hand, there is also a need to prepare tasks, processes and interactions in such a comprehensible way that employees are not only users but can also develop their own understanding of the processes taking place in the company. To address these aspects, model-based approaches have become established in science. This paper focuses on the model-based description and the interdependence between the PPC main tasks. The primary focus is on the PPC main task order management, which represents a cross-sectional task between individual PPC tasks, but also represents a continuous communication interface to the customer via order coordination tasks. In addition to these continuous functions in the order processing process, order management also has important tasks in the area of order clarification, feasibility checks and order acceptance, which are the focus of this article, especially for make-to-order (MTO) manufacturers (cf. [8]). The named tasks have an immense impact on the achievement of a company's objectives, as they control the load of the production.

In the literature, many approaches take a purely economic view (e.g. in the context of so-called revenue management [9]) and do not consider the complex interdependencies and dependencies with the other (main) tasks of PPC in sufficient detail. Therefore, this paper will also take a closer look at the strategic decisions that need to be made within order management and how these should be aligned with the information resulting from the other PPC tasks in order to optimally focus on the logistical objectives of companies. For this purpose, the essential basics of PPC and the main task of order management are first presented and then the interdependence between order management and other PPC tasks is discussed. A process view of the main task of order management derived from the Hanoverian Supply Chain Model (HaSupMo) is also presented, which addresses topics such as the strategic make-or-buy decision at the primary product level. The resulting conflicts of objectives in the context of order management are derived and presented transparently.

2. Fundamentals of PPC and order management

Taking a look at the tasks of PPC according to EVERSHEIM the planning and control of the manufacturing and assembly processes can be divided into schedule-related, capacity-related and volume-related tasks [10]. A central task, but one that is mostly underestimated, is order management. Because this is the interface to the customer, it is possible to influence the load on production directly but also indirectly. Thus, order management in the order processing process has the earliest and at the same time very strong possibility to control the load in production - e.g. through the strategic acceptance and rejection of customer orders - while at the same time also having a great influence on the profitability of the accepted orders.

Order management, therefore, has a significant influence on the achievement of the strategic corporate objectives of MTO manufacturers, which can be derived from the tactical and operational levels for the supply chain as an object of consideration. Looking at the main objectives of production logistics, high schedule reliability and short throughput times should be mentioned on the logistics performance side, but at the same time, the cost objectives of high capacity utilisation and low inventory levels should also be achieved [3,11,12].

A multitude of publications and frameworks can be found in literature (cf. [2]), dealing with the concept and the embedding of order management in corporate processes or PPC. The conceptual understanding of which tasks are to be located in order management strongly differs between the authors. Due to the large number of preliminary works, only the most important framework models and approaches for order management and

PPC will be mentioned here. For a more detailed overview and discussion of PPC frameworks respectively concepts such as Material Requirements Planning (MRP) and Manufacturing Planning and Control (MPC) as well as Hierarchical Production Planning (HPP) reference is made to MISSBAUER AND UZSOY [13].

2.1 Order management in existing PPC frameworks

A significant work in the field of order management is the one by WIENDAHL. Therefore, WIENDAHL defines the scope of application of order management following the Supply-chain operations reference (SCOR) model [14] as the procurement, production and delivery of goods, but differentiates order management from classic PPC systems. WIENDAHL criticises classic PPC systems, saying that there are deficits and too much complexity in the technology-oriented PPC systems (e.g. material requirements planning (MRP-) based) and develops a socio-technical, logistics-oriented approach to planning and controlling a supply chain that takes turbulence and disruptions into account. [15]

The extended Aachen PPC model is a comprehensive framework providing four perspectives on PPC. The described PPC tasks within the Aachen PPC model are differentiated into inter-company network tasks, internal core tasks and cross-sectional tasks. The main task of order management as a component of PPC is classified as a cross-sectional task and thus serves to integrate the network and core tasks and at the same time to optimise PPC. Order management further consists of the tasks of quotation processing, order processing and continuous integrated order coordination. [8]

The Hanoverian Supply Chain Model (HaSupMo) was first presented by SCHMIDT AND SCHÄFERS in 2017 [3]. The HaSupMo represents a synthesis of the considerations of the extended Aachen PPC model and the modelling approach of LÖDDING [16] and combines a PPC view, which differentiates main tasks and assignments, with a supply chain view, which structures the core processes along the company's internal supply chain with their specific logistical objectives. This enables a direct link between the PPC task and the logistical objective.

Within the model-based view of order management, an order passes through five tasks before real production orders are finalised and released (see Figure 1). After the task *clarify order* at the beginning, *roughly schedule production orders and plan safety time* as well as the customer order-related task *roughly plan resources* follow. Following these three tasks, the decisive check and subsequent decision are made as to whether this customer order can be realised.

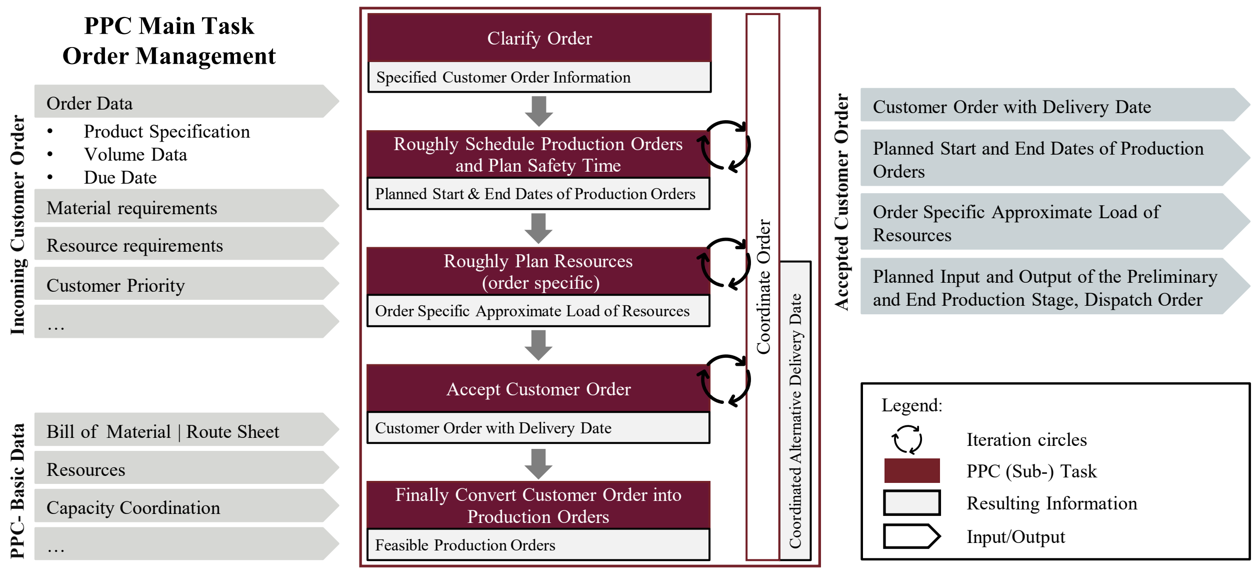


Figure 1: The PPC main task order management according to SCHMIDT [12]

If the check is positive, the order is accepted (*accept customer order*) and the subsequent task *finally convert customer order into production orders* takes place. Should the check be negative, however, the possible deviations of the delivery date are again coordinated with the customer (*coordinate order*). Only after renewed positive confirmation by the customer the two subsequent tasks of *accept customer order* and *finally convert customer order into production order*s take place [12,17]. In addition to the PPC tasks that run, Figure 1 also contains an overview of the possible incoming and outgoing information. For example, accepted sales orders contain information such as product specifications, quantity data and the customer's requested date, which is also referred to as the scheduled date. Basic data such as bills of materials (BOM), route sheets, resources or capacity coordination are already available in the PPC.

In a model-based view, accepted customer orders also generate a data output for PPC in form of the dispatch order and production program planning. Data such as contractually agreed delivery dates, planned start and end dates, planned resource allocation (planned capacity), but also planned inputs and outputs of the prepress and final stages of production and order dispatch are decisive objectives for further consideration in the HaSupMo.

2.2 Further research approaches in the field of order management

In addition to the two frameworks presented, there is a large number of works dealing with order management or with partial aspects or subtasks of order management.

For instance, approaches can often be found that deal with the topic of accept customer orders. The term revenue management [9,18], which has its origins in the airline and hotel business, is widespread and generally represents a variety of instruments for dealing with requests and limited production capacities. Other publications look at the question of how offers should be included in planning and how order acceptance and scheduling can be implemented, taking into account the acceptance rate [4,19]. Subjects such as combined production planning and contractors and demand uncertainty are also addressed [20]. KATE [21] describes two extreme examples in this context with regard to the design of order taking, which differ in terms of their degree of integration (cf. [22]), which describes the system characteristics between a completely hierarchical and a completely integrated planning procedure, of the Sales and Production areas.

Other approaches deal with decisions for or against the outsourcing of orders to suppliers or, in other words, the make-or-buy decision for primary requirements (finished goods) [23] or also the topic of the strategic inclusion of products from outside the portfolio to utilise production capacity [24]. There are also OR models for the integration of outsourcing decisions into PPC (cf. [25]). A good overview of approaches and decision determinants in this context was presented by SERRANO [26].

2.3 Interim conclusion

It is shown that an integrative view of order management with the other main PPC tasks enables possibilities for the comprehensive and target-oriented derivation of both order acceptance as well as outsourcing decisions. However, it was identified through a literature review that previous frameworks for PPC deal with the procedures and tasks of PPC in isolation and (deliberately) greatly simplify the complexity that would arise in particular with an integrative view across the boundaries of tasks.

This paper, therefore, builds on the existing frameworks and approaches and derives an extended representation of the task at hand and essential decision-making processes in order management, taking into account the interdependencies with other main PPC tasks and addressing conflict of objectives. The modelling is carried out in accordance with the previous descriptions in the HaSupMo - serving as a suitable modelling basis - and can thus be seen as an extended view, which transforms and supplements the existing knowledge of the HaSupMo from the previous task-related view into a more time sequence-oriented view. Building on the expanded view of order management created in this way, which also considers in particular the make-or-buy decisions and the question of order acceptance in companies, the resulting conflicts of objectives are discussed.

3. An integrated process model of order management

Between the individual tasks of PPC there is a multitude of interdependencies and dependencies, which have so far mostly only been investigated on a qualitative level with a very limited focus of observation (cf. [2]). In order to approach the question of the existing interdependencies between order management and the other main PPC tasks and the associated conflicts of objectives, an extended process-oriented view of the task of order management in the modelling logic of the HaSupMo was created in condensed form. Figure 2 shows the derived extended (process-oriented) view of the main PPC task of order management.

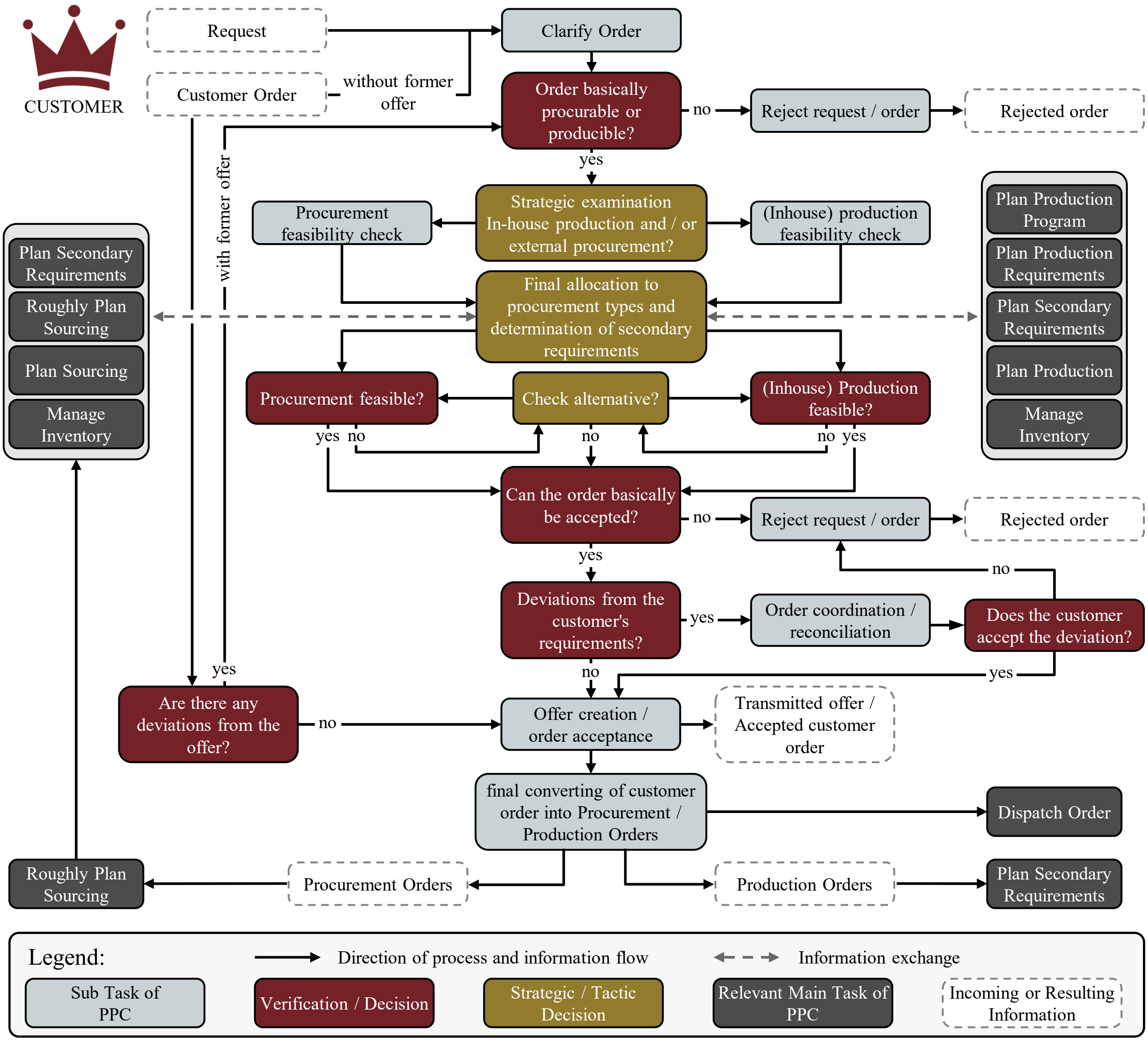


Figure 2: Extended process-oriented view of the task of order management in the modelling logic of the Hanoverian Supply Chain Model according to SCHMIDT AND SCHÄFERS [17] and SCHMIDT [12]

The main task order management is initiated by a customer request or an incoming customer order. If it is a request or a direct customer order for which there was no previous quotation, the first step is to *clarify the order*, which leads to a decision as to whether the customer's needs can in principle be met by the company. Within the framework of this inspection, organisational restrictions are not yet included, rather it is about the possibility of technical implementation, which is not given, for example, if the corresponding machines are not available and the order cannot be subcontracted to suppliers. If this first check is negative, the order or request must be rejected; if it is positive, a strategic decision must be made as to whether the request or order should be produced in-house or outsourced (*make-or-buy decision*). The strategic decision at this point focuses on the production of the primary requirement, e.g. the fully assembled article, and is usually made taking into account a large number of influencing variables and conflicting objectives. For example, the decision to outsource should not lead to the company's own production being drastically underutilised, which would also reduce its profitability.

In order to be able to derive a comprehensive decision, information is needed that is provided by other main PPC tasks. For example, the principle delivery capability of suppliers (resulting from the PPC main task *roughly plan sourcing*) must be checked as well as an overview of the utilisation of production that is as exact as possible, which, depending on the time horizon, includes the result of the PPC main tasks *plan production program*, *plan production requirements* or *plan production* in the considerations.

When the procurement and/or in-house production of the primary requirement has been evaluated, the procurement type must be finally determined and then the pourability or producibility of the dependent requirements (e.g. assemblies) must be checked. This requires the information from the task *plan secondary requirements*. Particularly at the level of secondary requirements, it may be possible that for some products both in-house production and external procurement are possible, which makes it possible, for example, to relieve the burden on in-house production through targeted external procurement. In order to carry out such a check, information is needed from the main PPC tasks of *external procurement* and *in-house production* as well as strategic decisions regarding questions such as know-how transfer to the supplier.

As soon as the procurement type for all secondary requirements has been carried out, the customer's request / order can be answered. Depending on the results of the previously performed checks, deviations from the customer's requirements, e.g. with regard to materials or the schedule, may result and need to be communicated. If the customer accepts the adjustments or if the request / order can be confirmed unanimously, the offer is sent to the customer or the acceptance of the order is confirmed. If the customer places an order on the basis of an existing quotation and no changes have been made to the quotation, the latter shall also take place.

As a final step, the customer order is converted into procurement and/or production orders and further processing is initiated in the other main PPC tasks.

4. Conflicting objectives and interdependence in order management

As already indicated in section 3, direct and indirect (reciprocal) interactions exist between the PPC main task order management and the other PPC main tasks. In the adjacent production program planning, the exchange of information with order management is important to integrate the accepted customer orders into the existing production program. At the same time, there is a close data exchange between order management and resource planning dependent on it. Secondary requirement planning, therefore, converts the information received from order management (mainly about the accepted production orders) and conducts a make-or-buy decision on the basis of this information. Depending on the type of procurement of primary and/or secondary requirements, there are various effects on the logistical objectives of a company. Decisions for or against outsourcing influence various control variables in production. These include the planned output and the planned input, whose difference determines the inventory level of production and thus has an impact on the objective work-in-process. This in turn significantly determines the realisable throughput times and thus also the extent to which the customers' target specifications can be met. If the inventory level becomes too high, for example because the input outweighs the output, it is necessary to initiate compensatory measures in order to be able to meet the specified delivery dates. Since order management has a striking influence on the initial load and thus the input of a production, it follows that reactions may become necessary within the framework of the PPC tasks plan production program and plan secondary requirements and that there is an immense influence on the achievement of logistical objectives.

If these interdependencies, which initially follow the order management in terms of time, are considered during the order management tasks, various conflicting objectives must already be taken into account. Here, economic or cost-oriented objectives (e.g. high capacity utilisation and high throughput) and logistics performance objectives (e.g. low throughput times, high schedule reliability) are opposed to each other.

Taking a look at the extended process-oriented view of order management (Figure 2), the consideration of the make-or-buy decision for primary and secondary requirements contained therein can have an immense influence on the conflicting objectives. Especially through the integrated view within order management, positive effects of the compensation measures on the achievement of objectives in production logistics can be expected, which can be exemplified in the following three cases.

Case 1 (Overload)

An MTO manufacturer whose production is already fully utilised in the current and future planning period receives an additional order from a strategic important customer at short-term notice. The new order is to be assigned to in-house production after inspection. According to the assessment of production utilisation, the additional order acceptance would lead to production overload. Usually, this results in an undesired order rejection or delivery date deviation. In contrast, if current independently produced primary or secondary requirements can also be assigned to external procurement, active control of the company's own production capacity utilisation is already possible at an early stage through order management. In this case, not only is it possible to accept orders in short term, but the decision to outsource also simultaneously creates the possibility of optimising company-internal economic and logistical objectives. Furthermore, accepting orders in short term results in higher turnover and at the same time avoids overloading production. Nevertheless, the objective of high schedule reliability remains and is not directly affected.

Case 2 (Underload)

In comparison to Case 1, the situation differs for companies that already procure a certain proportion of their primary and secondary requirements from external sources for reasons of capacity or economic efficiency. In case of an unforeseeable decrease in demand due to a (significant) reduction in customer requests, the consequence would be that the company's own production capacities for in-house production would only be used to a limited extent since a lower production volume would also be called up there. Using the early make-or-buy decision in order management, the products that are suitable for both in-house production and external procurement can now be produced in-house in short term by converting from buy to make. As a result, this compensatory measure can also maintain the company's internal target of production utilisation from in-house production at a similar initial level despite the declining turnover.

Case 3 (Order fluctuations and uncertainties)

If companies are generally exposed to many fluctuations or uncertainties regarding the order situation, the make-or-buy decision within the framework of order management enables them to cope with the constantly changing production load and thus with the utilisation. The order-related allocation of primary and secondary requirements to external procurement or in-house production enables access to capacities that can be used strategically. With the extended process-oriented view of the task of order management (Figure 2), both a short-term increase and a decrease in capacity are possible. A short-term capacity increase for the acceptance of additional customer orders for primary requirements is possible if the acceptance of additional customer orders allocates the new requirements to external procurement. At the same time, the economic objective of turnover is increased. Nevertheless, this approach is mostly associated with a higher risk and more extensive planning. The model can also be applied in reverse. This type of application should be chosen if a capacity reduction is necessary for economic reasons. In this case, the new requirements of the accepted customer orders must be allocated to in-house production at an early stage in order to keep production utilisation as constant as possible.

Resume

The presented form of order-related capacity flexibility brings significant advantages. Short-term fluctuations in demand can be absorbed while adhering to the company's internal targets. Active control thus enables the right steps to be taken at an early stage in order to survive on the market in the long term.

Even a strategic shift of orders or individual requirements to suppliers is a thinkable option in order to accept further, possibly even more lucrative orders. However, the economic efficiency and strategic importance of outsourcing should be examined in advance in all respects.

Apart from that, it is important to check which parts are suitable for both external procurement and in-house production. After this verification and selection of these parts, further important factors have to be taken into account for a target-oriented application of the described advantages. On the one hand, there are hard factors that define the framework of sourcing possibilities through legal and contractual requirements. These include, for example, the certification of suppliers, internal audits to ensure quality and working methods, political environmental targets or contractual requirements from customers. On the other hand, there are also soft factors that take into account the company's internal strategic actions. These include, for example, key performance indicators (KPI´s) and logistical objectives, knowledge transfer to other suppliers, the classification of suppliers from previous orders and supplier management to safeguard the company's own production.

5. Conclusion and Outlook

In this paper, it was shown that order management as a PPC main task with its input and output data has a significant influence on the logistical objectives of company-internal supply chains for MTO manufacturers. With the extended process-oriented view of the PPC main task order management, it was shown that the active control of primary or secondary requirements using a target-oriented make-or-buy decision can positively influence the production load and utilisation. This form of integrated capacity flexibility can thus be used for the targeted levelling of load fluctuations if products / parts can be assigned to both in-house production and external procurement. Building on this, economic and logistical trade-offs were identified from which strategic and economic decision-making options for MTO manufacturers to optimise their own production capacity utilisation can be derived.

Further research activities should focus on the interrelationships and effects of selected order management procedures on the logistical objectives. In particular, the minimisation of conflicts between logistical objectives through selected procedures should be considered. Furthermore, this more detailed investigation could lead to an optimisation approach of the strategic decision for internal supply chains of MTO manufacturers with regard to outsourcing in the short and medium-term. In addition, further consideration should be given to capacity expansion or minimization enabled by outsourcing opportunities to optimize internal company goals.

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] Schuh, G., 2007. Effizient, schnell und erfolgreich: Strategien im Maschinen- und Anlagenbau. VDMA-Verl., Frankfurt, M., 94 pp.

[2] Mütze, A., Hillnhagen, S., Schäfers, P., Schmidt, M., Nyhuis, P., 2020. Why a systematic Investigation of Production Planning and Control Procedures is needed for the target-oriented Configuration of PPC, in: 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 103–107.

[3] 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.

[4] Mundt, C., Lödding, H., 2020. Order Acceptance and Scheduling with a Throughput Diagram, in: Advances in Production Management Systems. The Path to Digital Transformation and Innovation of Production Management Systems. IFIP WG 5.7 International Conference, APMS 2020, Novi Sad, Serbia, August 30 – September 3, 2020, Proceedings, Part I, Cham. 2020. Springer International Publishing, Cham, pp. 351–359.

[5] Seitz, M., Härtel, L., Hübner, M., Merkel, L., be Isa, J., Engehausen, F., Meluzov, N., Rost, R., Schmidhuber, M., Sauermann, F., Hünnekes, P., 2018. PPS-Report 2017/18: Studienergebnisse, 1. ed. TEWISS, Garbsen, 42pp.

[6] 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: Studienergebnisse, 1. ed. TEWISS, Garbsen, 50 pp.

[7] INFORM, 2020. Trendreport - Der Maschinen- und Anlagenbau 2020: Eine zukunftssichere Branche in Deutschland?: INFORM Institut für Operations Research und Management GmbH Juni 2020. https://www.inform-software.de/produkte/felios/trendreport-maschinenbau. Accessed 17 December 2020.

[8] Schuh, G., Stich, V., 2012. Produktionsplanung und -steuerung 1: Grundlagen der PPS, 4. Aufl. 2012 ed. Springer Berlin Heidelberg, Berlin, Heidelberg.

[9] Defregger, F., Kuhn, H., 2004. Revenue management in manufacturing, in: Operations research proceedings 2003 : selected papers of the International Conference on Operations Research (OR 2003), Heidelberg, September 3 - 5, 2003 ; with 51 tables. Springer, Berlin, pp. 17–22.

[10] Eversheim, W., 2002. Organisation in der Produktionstechnik 3: Arbeitsvorbereitung, 4., bearbeitete und korrigierte Auflage ed. Springer, Berlin, Heidelberg, 295 pp.

[11] Nyhuis, P., Wiendahl, H.-P., 2009. Fundamentals of Production Logistics: Theory, Tools and Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, 320 pp.

[12] Schmidt, M., 2018. Beeinflussung logistischer Zielgrößen in der unternehmensinternen Lieferkette durch die Produktionsplanung und -steuerung und das Produktionscontrolling. Habilitationsschrift, Hannover.

[13] Missbauer, H., Uzsoy, R., 2020. Production planning with capacitated resources and congestion. Springer, New York NY, 285 pp.

[14] Huan, S.H., Sheoran, S.K., Wang, G., 2004. A review and analysis of supply chain operations reference (SCOR) model. Supply Chain Management: An International Journal 9 (1), 23–29.

[15] Wiendahl, H.-H., 2011. Auftragsmanagement der industriellen Produktion. Springer Berlin Heidelberg, Berlin, Heidelberg, 474 pp.

[16] Lödding, H., 2013. Handbook of manufacturing control: Fundamentals, description, configuration. Springer, Heidelberg, 577 pp.

[17] Schmidt, M., Schäfers, P. The Hannoverian Supply Chain Model (HaSupMo): Interactive Hompage. Institute of Production Systems and Logistics. http://hasupmo.education/. Accessed 20 June 2020.

[18] Rehkopf, S., 2006. Revenue Management-Konzepte zur Auftragsannahme bei kundenindividueller Produktion. DUV, Wiesbaden.

[19] 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 (2), 196–209.

[20] Aouam, T., Geryl, K., Kumar, K., Brahimi, N., 2018. Production planning with order acceptance and demand uncertainty. Computers & Operations Research 91, 145–159.

[21] Kate, H.A. ten, 1994. Towards a better understanding of order acceptance. International Journal of Production Economics 37 (1), 139–152.

[22] Steven, M., 1994. Hierarchische Produktionsplanung, 2., überarbeitete und erweiterte Auflage ed. Physica-Verlag HD; Imprint; Physica, Heidelberg.

[23] Chen, K., Xiao, T., 2015. Outsourcing strategy and production disruption of supply chain with demand and capacity allocation uncertainties. International Journal of Production Economics 170, 243–257.

[24] May, M.C., Schmidt, S., Kuhnle, A., Stricker, N., Lanza, G., 2021. Product Generation Module: Automated Production Planning for optimized workload and increased efficiency in Matrix Production Systems. Procedia CIRP 96, 45–50.

[25] Lee, Y.H., Jeong, C.S., Moon, C., 2002. Advanced planning and scheduling with outsourcing in manufacturing supply chain. Computers & Industrial Engineering 43 (1-2), 351–374.

[26] Serrano, R.M., Ramírez, M.R.G., Gascó, J.L., 2018. Should we make or buy? An update and review. European Research on Management and Business Economics 24 (3), 137–148.

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.

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 Hanover 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.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Collaborative Factory Layout Planning With Building Information Modeling

Thomas Neuhäuser[1,2], Philipp Michaeli[3], Lisa Lenz[4], Andrea Hohmann[1], Fabian Matschinsky[5], Franz Madl[6], Rayk Henkelmann[7]

[1]*Fraunhofer Institute for Casting, Composite and Processing Technology*
[2]*Institute for Machine Tools and Industrial Management, Technical University of Munich*
[3]*ifp - Prof. Dr.-Ing. Joachim Milberg Institut für Produktion und Logistik GmbH & Co. KG*
[4]*Department of Construction Management, TU Dortmund University*
[5]*HL-Technik Engineering GmbH*
[6]*pbb Planung + Projektsteuerung GmbH*
[7]*HIB – Henkelmann Ingenieur Büro GmbH*

Abstract

Manufacturing companies are facing an increasingly turbulent environment. 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 the lack of interoperability. Therefore, new ways have to be found in factory planning to overcome these problems. Building Information Modeling (BIM), which is used in the construction industry, provides a promising method for collaboration and interaction in complex projects. Therefore, an approach is presented to systematically implement the BIM method in the factory layout planning process. The aim of this approach is to achieve a higher degree of collaboration between the individual specialist planners in earlier planning phases in order to reduce time and costs over the entire project and to lead to better and more valid planning decisions and results. For this purpose, the individual roles in the planning process are defined, a process diagram for layout planning is drawn up, and a basic data set is specified for the systematic entry of the trades.

Keywords

Factory Planning; Building Information Modeling; Layout Planning

1. Introduction

The manufacturing industry is one of Germany's most important economic sectors, accounting for 22.9% of gross domestic product [1]. However, companies in the economic sector are facing challenges and trends such as globalization, dynamization of product life cycles and climate change. These circumstances lead to a turbulent market environment for manufacturing companies. Therefore, innovations are developed more regularly, to which the production system has to be adapted within the shortest possible time. Production adjustments are hence becoming more frequent and must be implemented much faster. Factory planning thus becomes a continuous task. [2–4]

The central target variables of planning projects are time, costs and quality. While the quality targets are generally met, the time targets are missed in about 60% of projects and the cost targets in approximately 72%. This is partly due to a lack of interoperability between the different planning participants [5]. A study by Gallagher et al. [6] concludes that the lack of interoperability in the facilities industry costs $15.8 billion per year in the U.S. alone. To overcome these problems, new planning approaches must be developed and applied in factory planning projects.

The methodology of BIM starts exactly at this point. BIM is a collaborative planning method that allows the systematic exchange of data and information between stakeholders throughout the entire lifecycle of a building based on digital models, which contain geometrical data as well as non-geometrical data like floor loadings. With an agile planning approach and standardized exchange formats in the openBIM process, the lack of interoperability can be avoided and processes are more efficient. Current studies show that the use of BIM in planning projects can reduce time and costs by up to 25%. [7–9]

While BIM has been mandatory for public infrastructure projects in Germany since 2020, there have been few known use cases in factory planning [10]. The main reasons for this are the lack of support for cross-discipline data exchange formats and the systematic separation of production system and building planning procedures [9]. To overcome these shortcomings, this paper introduces a BIM-based collaborative way of working in factory planning projects. 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. For this purpose, chapter 2 explains the fundamental differences between classical and BIM-based planning procedures. Chapter 3 sets out and subsequently analyzes the information exchange requirements between the single specialist planners. In chapter 4 the implications for the factory layout planning process are discussed. Finally, the paper ends with a summary and a conclusion.

2. Fundamentals

2.1 Conventional planning procedure

Factory planning is composed of the planning of the production system, which is responsible for the value creation of the company, and the building, which forms the envelope around the production. The VDI guideline 5200 part 1 [11] describes the classical planning approach of the production system. According to this, the planning process is structured in seven successive phases, ranging from phase 1 setting of objectives to phase 7 ramp-up support. The “Honorarordnung für Architekten und Ingenieure” (HOAI 2021) [12] describes the planning procedure for construction projects in Germany. Therein, nine different service phases are passed through from phase 1 the establishment of the product basis to phase 9 the project management and documentation. Figure 1 shows the assignment of the phases of production system and building planning to the factory planning process. [11,12]

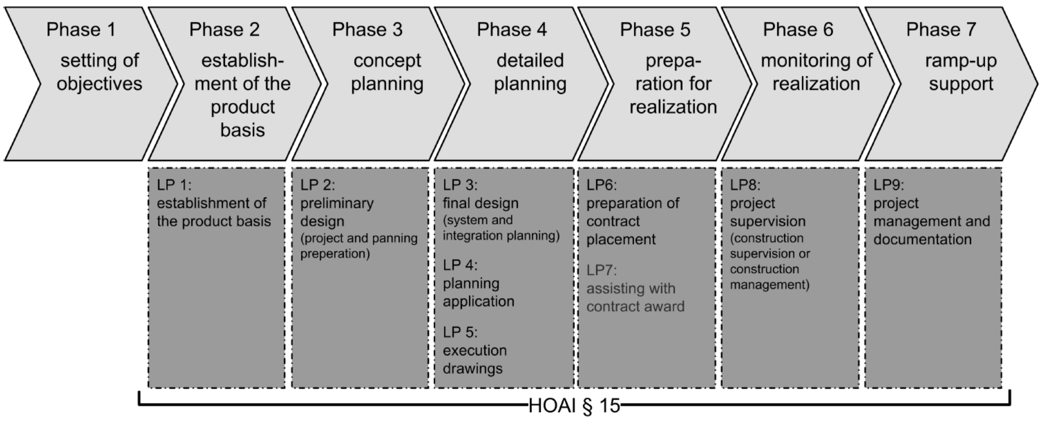


Figure 1: Assignment of performance phases according to HOAI, Article 15, to the planning phases [11]

Layout planning takes place in phases 3 and 4 of the production system planning process. It is divided into three steps, ideal planning, real planning and fine planning. The input parameters for ideal planning are the dimensioned resources, a rough logistics concept and a scale function diagram showing space assignments. In the ideal planning itself, structural units such as resources are then ideally arranged in relation to one another. Several layout variants can be designed, for each of which a building envelope is designed. In the real planning, the layout variants are merged with the building envelope and the structural framework is designed. Subsequently, the generated variants are evaluated monetarily, quantitatively and qualitatively and a feasible factory concept is decided upon. In the subsequent fine planning, the material, information and communication flows, the equipment, the work organization and the media supply concept are determined. This results in a fine layout with specified operating resources and a complete building design, which can then be used to prepare the permit applications. The phases of layout planning of the production system correspond to the phases 2 and 3 of the HOAI 2021. [11]

However, the systematic separation of production system planning and building planning in two different directives, respectively regulations, often leads to a strict separation of the two in projects. In particular, short-term adjustments in both disciplines lead to misunderstandings, planning errors, uncoordinated processes, delayed results as well as non-compatible data formats and thus to a failure to meet time and cost targets in factory planning projects.

2.2 BIM-based planning procedure

BIM-based planning approaches attempt to overcome these problems by systematically connecting the individual planning disciplines, by bringing forward planning and decision-making processes, as shown in figure 2, and by establishing the standardized data exchange format Industry Foundation Classes (IFC) in the openBIM process. Over 300 different software tools for various specialist planners already support IFC, which greatly simplifies the exchange of data and information and counteracts the lack of interoperability [13]. In addition to IFC, a second exchange format is relevant for a BIM-supported planning process, the so-called BIM collaboration format (BCF) [14]. This vendor-neutral format is used to exchange coordination messages in change management of a design process. If, for example, a collision occurs between different specialist models, a report with the position, perspective, affected object and text can be distributed to the specialist planners. [7]

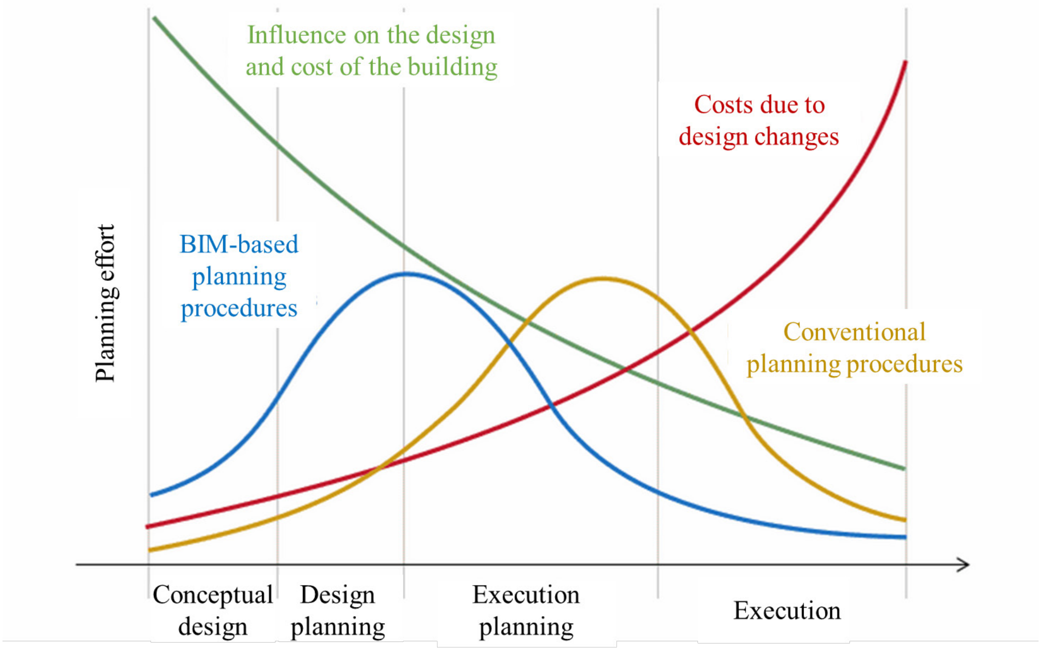


Figure 2: Forward shifting of planning and decision making processes in BIM-based planning procedures [7,15]

In addition to the data exchange formats, the individual roles in the planning process are also crucial. Additionally, to the individual specialist planners, a so-called BIM manager is introduced, who is responsible for several tasks, e.g. to take over quality assurance in the project. To this end, the BIM manager regularly brings together the isolated planning statuses of the specialist planners and checks them for clashes. The specialist planners themselves are divided into authors and coordinators of their subject-specific BIM model. The authors provide the content for the specialized model. The coordinator of a specialist discipline is responsible for the provision and the quality of the developed model. The authors and coordinators thus always cover a specialist discipline, while the BIM manager should be independent of the discipline. [7,14]

To implement a BIM-based approach in a project, the BIM targets and use cases must be defined after setting the project and factory objectives. The concrete implementation of the BIM methodology in the project is then derived from these. It is advisable to create an Information Delivery Manual (IDM) in order to define which planning participants have to transfer which data to whom and at what time. These data delivery points are referred to as data drops [16]. However, such an IDM is not yet available for general layout planning process and is therefore described in the following chapter.

3. IDM for the BIM use case factory layout planning

To close this gap and generating the IDM for the use case of factory layout planning, an inductive research approach was chosen. In several workshops with factory planning experts from industry and research, it was defined which planning participants have to transfer which data to whom and at what time.

In the presented approach, only the layout of the production facilities is considered in layout planning. Influencing factors such as the logistics system, work aids or even the people in the factory are not taken into account. Moreover, no distinction is made between BIM authors and coordinators, but only the role of the specialist planning discipline as a whole is considered.

3.1 Description of roles

As shown in table 1, there are four basic planning disciplines and the BIM manager in the factory layout planning process. The disciplines are the production system planner, who plans the layout, the architect, who designs the building envelope, the structural engineer, who plans the structural framework, and the building equipment (MEP) planner, who plans e.g. the media supply and disposal.

Table 1: Description of roles in the factory layout planning process

Role	Description
BIM manager	Coordination of the individual specialist planners, ensuring quality of the planning process and the absence of clashes as well as control of change management
Production system planner	Plans the layout from ideal to real to fine
Architect	Plans the building envelope
Structural engineer	Plans the structural framework including the foundations
MEP planner	Plans the media supply system including the IT connection as well as electrical engineering and the discharge of emitted media around the production facilities

3.2 Process diagram

The stakeholders intervene in the planning process at different points in time and develop its contents. Figure 3, 4 and 5 show the interlocking as Business Process Model and Notation diagrams throughout the layout planning phases.

The process begins in the ideal planning phase (figure 3) with the BIM manager. The manager creates a coordination body, as an empty coordination model with a coordinate system. This ensures that each discipline model is planned in a uniform coordinate system. This coordination model is passed on to the production system planner as an IFC file. The latter carries out the ideal layout planning and thus generates the first model of the production system (Production system model_v1). In contrast to the theoretical approach, in which several variants are created, often only one ideal variant is generated. This is used again and again as an ideal benchmark in the further course of planning.

Subsequently, the architect begins to define the building restrictions. The models are always transferred as IFC files and normally there are no iterations between the production system planner and the architect. Afterwards, the next planning phase begins.

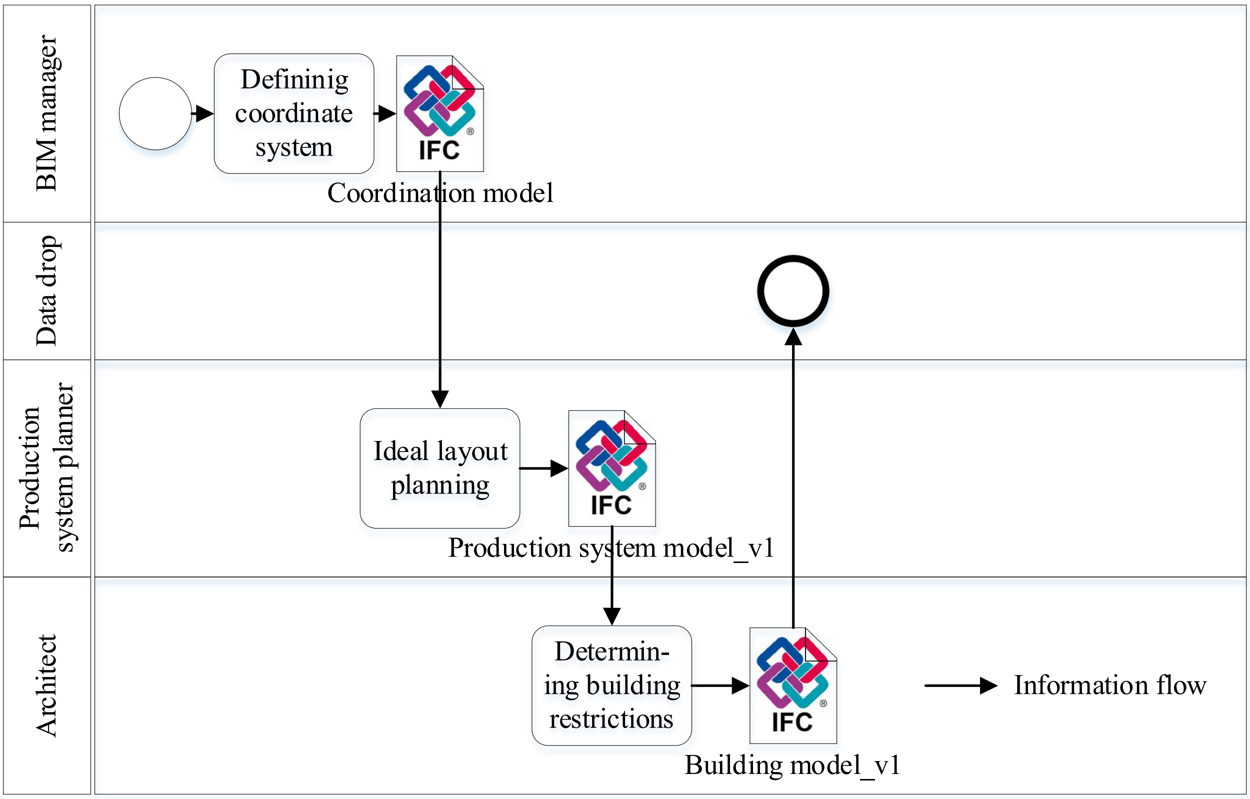


Figure 3: Process diagram for the ideal planning

The structural engineer and building services engineer are involved in the real planning (figure 4). At the beginning of this phase, the production system planner creates several real layout variants based on the ideal layout and the building restrictions. These variants can follow different target points such as expandability, material flow optimization or lead times. These variants are then roughly planned and evaluated by the other planning disciplines. It must be ensured that a relatively high level of abstraction is still maintained, but that the individual variants can still be validly evaluated by the individual specialist planners. Even in this phase, it is not necessary to follow a purely sequential process; instead, there may be several iterations.

This phase is followed by a clash detection, after the final data delivery as a big data drop. It is crucial that the models are collision-free and that the project and factory objectives are met. The BIM manager performs these tasks, for both the production system as well as the building planning. Adherence to the objectives is decisive in that the degrees of freedom of the planning and thus the specified restrictions must be moderated between the production system planner and the architect. For example, an objective such as the transformability of the building must not be allowed to get out of hand when planning should be optimized for material flow. The moderation of the focus is decisive because, on the one hand, the building can be expected to have a significantly longer service life than the production system and its adaptability can therefore be very important. On the other hand, an actual greenfield must not become a brownfield planning due to overly tight building restrictions, as a result of which planning close to the ideal layout is no longer possible. To ensure that there are no clashes, the BIM manager must merge the individual models in a collision check. If clashes are detected, change orders are distributed to the specialist planners as BCF files.

If there are no collisions or if all collisions are eliminated, a decision will be made for one of the variants based on the evaluation of the individual planners (Factory model_v1). The ideal variant is used as a benchmark and the evaluation takes place quantitatively, monetarily and qualitatively. Afterwards, the fine planning begins.

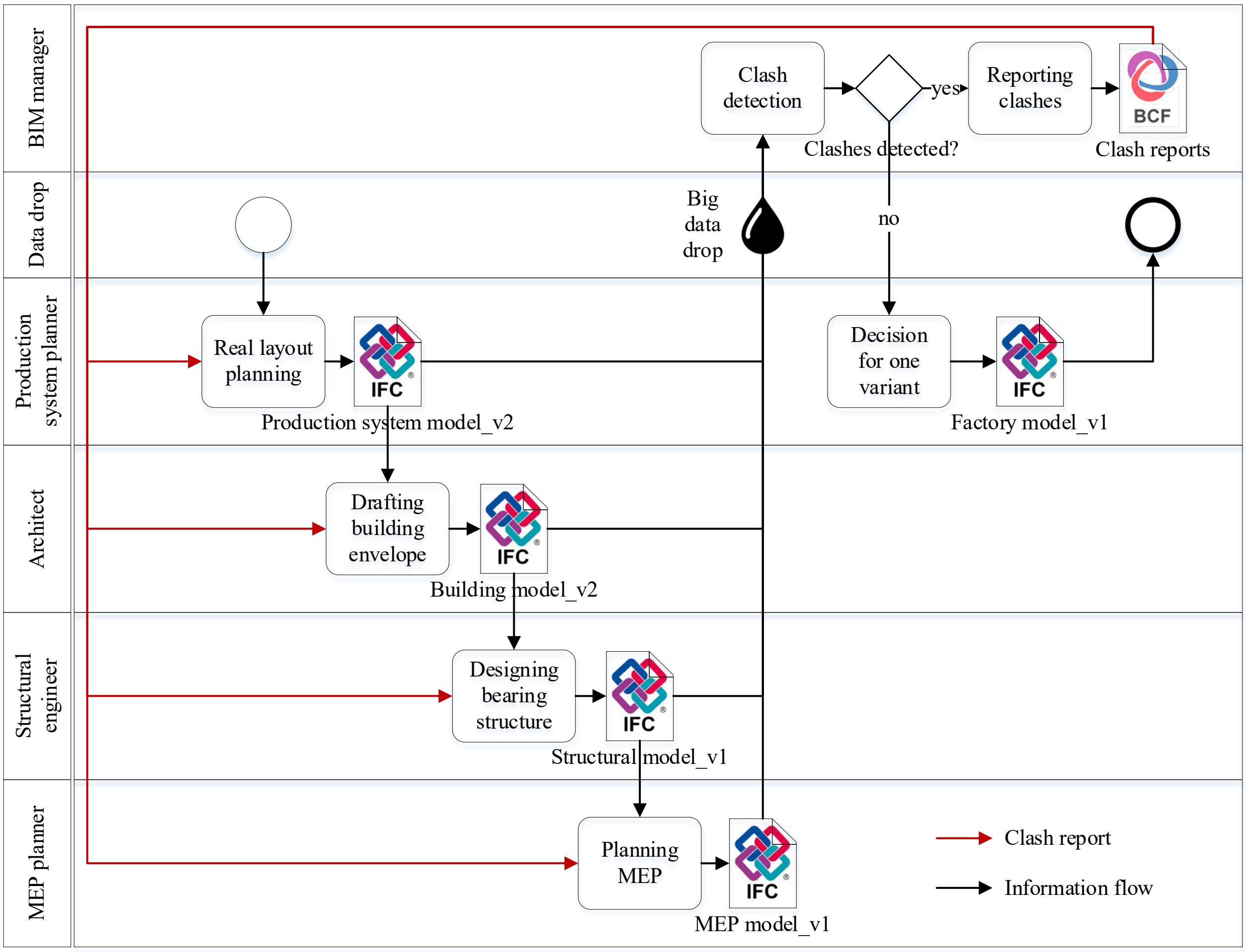


Figure 4: Process diagram for the real planning

In fine planning as in real planning, all planning disciplines are involved (figure 5). Here, there can be significantly more iteration loops, since the high level of detail in planning means that significantly more data and information must be exchanged. So the final big data drop at the end of the phase is joined by many small data drops in the iteration loops. These are delivered every one to six weeks, depending on the agreed interval duration, also called sprints agile planning approaches. Once the collision check has been completed, the factory model (Factory model_v2) is available and the permit applications can be prepared.

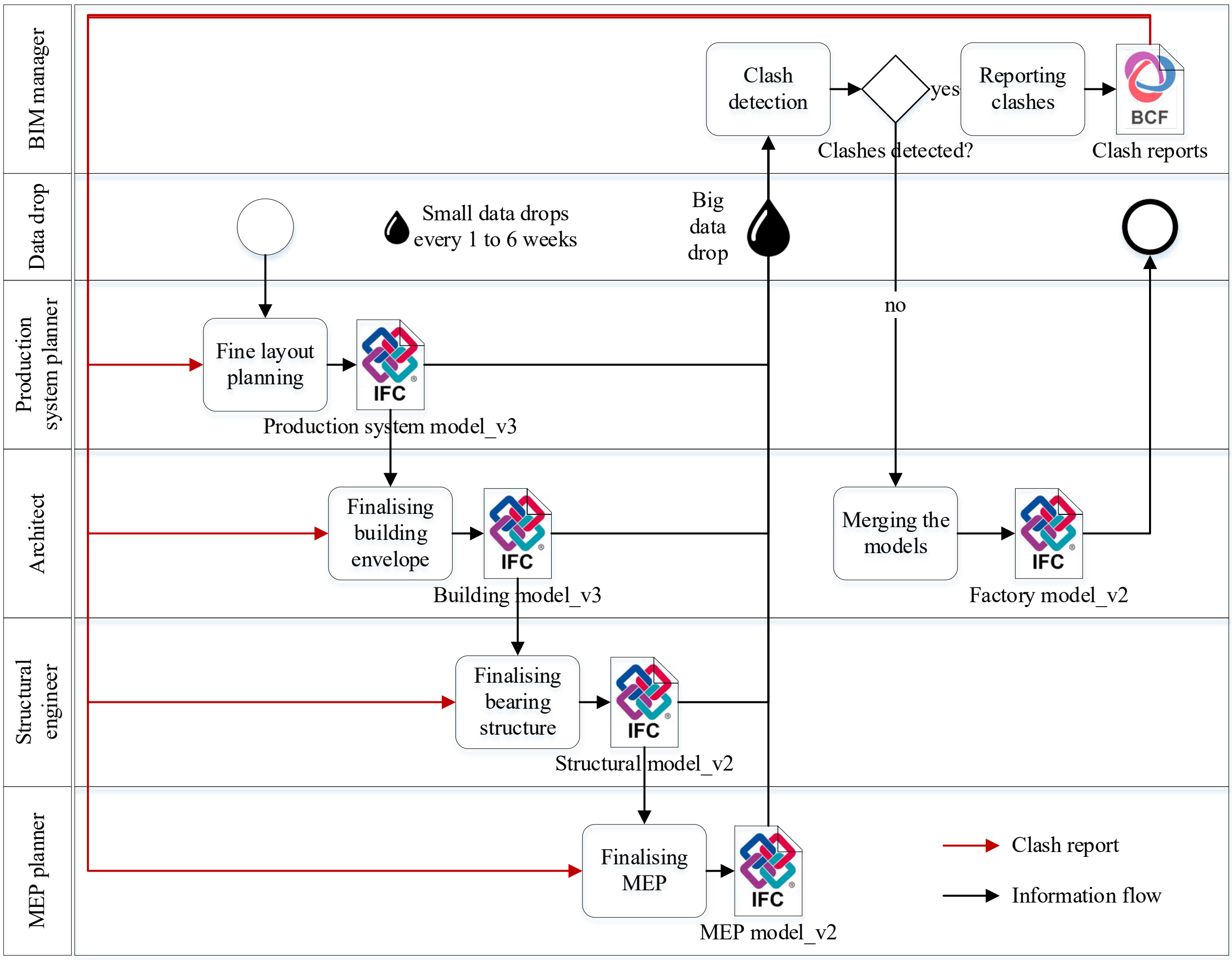


Figure 5: Process diagram for the fine planning

3.3 Exchange Requirements

In the case of exchange requirements, the second production system model (Production system model_v2) is of particular importance. In a BIM-based planning procedure, this must already contain a first data set on which the other specialist planners can carry out their initial rough planning. This information, which is all related to the production facilities, includes the spatial arrangement, the geometric dimensions, the weight, the vibration behavior, the media requirements and connections, the electrical power requirements and connections, the emitted media quantities and their connection points, certain safety as well as fire protection requirements and IT connections. The information types as well as the required information are shown in detail in table 2.

Table 2: Exchange Requirements for the second production system model

Type of information	Information needed
Spatial arrangement	Spatial arrangement in relation to other production facilities
Geometric dimensions	Description of outer contours
Weight	Weight
Vibration behavior	Description of frequency and amplitude of the vibration behavior
Media requirements	All operating media such as cooling lubricants, process gases, compressed air or even water according to type and quantity
Media connections	Place and type of connection
Electrical power requirements	Type and quantity of electrical demand
Electrical power connections	Place and type of connection
Emitted media quantities	All emitted media such as exhaust air, heat, used cooling lubricants or also waste water according to type and quantity
Emitted media connections	Place and type of connection
Safety requirements	All safety-related requirements such as protective fences and safety zones
Fire protection requirements	All fire protection-related requirements like sprinklers
IT connections	Place and type of connection

This data set does not have to be complete and may contain some uncertainties, but should be updated and exchanged throughout the planning process. The quality assurance of this process can be ensured by the BIM manager through regular quality gates.

4. Implications for the factory planning process and discussion

A BIM-based factory planning approach leads as shown in chapter 3 to a significant increase in planning efforts in the early phases. For example, the production system planner must provide the most detailed information possible on the production facilities as early as the second model so that the other specialist planners can build on this. In addition, the technical building equipment planner is involved in the planning process at a much earlier stage and has to evaluate planning alternatives for different layout variants. This definitely leads to an increased planning effort, which, however, can make the subsequent phases less time-consuming and lead to better planning results due to the clearly well-founded planning. In addition, this leads to a more valid cost estimation for the client.

If the opportunity for better planning results is set against the risk of increased planning effort in terms of time and costs, the risk should not be shied away from. Only 1% of the costs incurred in the life cycle of a structure are attributable to the planning phase, whereas 90% are incurred in the operating phase. The remaining 9% is accounted for by implementation. Accordingly, a trade-off between a reduced planning effort and a worse planning result is only worthwhile from a factor of 90. [17]

5. Summary and conclusion

In the paper, an approach for a BIM-based factory planning procedure was presented. Based on the role definitions in the planning process, a process diagram for factory layout planning was developed and the information exchange requirements for the first production system model were defined. With the help of this model, the other disciplines can be integrated into the planning process much earlier. This early involvement can result in additional costs in the early planning phases, which can be more than offset by a reduction in the workload in the later phases or a reduction in costs in factory operation and the whole life cycle of a building. In addition, BIM has the potential to support the achievement of cost and time targets in factory planning projects by the systematic integration of production system and building planning processes, and the standardized exchange of information can significantly reduce costs due to a lack of interoperability.

The next step is to complete the information exchange requirements for all handover documents to generate a complete IDM for the factory layout planning process. The Model View Definitions can then be derived from this and an IFC extension can be made to include production system data. This is currently being developed within the framework of *VDI Guideline 2552 part 11.8 BIM – factory planning* and the *buildingSMART Roundtable openBIM in factory planning*. In addition, the approach will be applied in pilot projects to analyze the implications in reality.

References

[1] Statista, 2021. Anteil der Wirtschaftszweige an der Bruttowertschöpfung in Deutschland im Jahr 2020.

[2] Abele, E., Reinhart, G., 2011. Zukunft der Produktion: Herausforderungen, Forschungsfelder, Chancen. Carl Hanser Fachbuchverlag, s.l., 262 pp.

[3] Pawellek, G., 2014. Ganzheitliche Fabrikplanung: Grundlagen, Vorgehensweise, EDV-Unterstützung, 2. Aufl. ed. Springer Vieweg, Berlin, 443 pp.

[4] Wiendahl, H.-P., Reichardt, J., Nyhuis, P., 2014. Handbuch Fabrikplanung. Carl Hanser Verlag GmbH & Co. KG, München.

[5] Reinema, C., Pompe, A., Nyhuis, P., 2013. Agiles Projektmanagement. ZWF 108 (3), 113–117.

[6] M. Gallaher, A. O'Connor, John L. Dettbarn, L. T. Gilday, 2004. Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry, in: .

[7] Borrmann, A., König, M., Koch, C., Beetz, J. (Eds.), 2015. Building Information Modeling: Technologische Grundlagen und industrielle Praxis. Springer Vieweg, Wiesbaden, 591 pp.

[8] Dr. Wieselhuber & Partner GmbH, 2018. BIM – Are you Ready?: Strategische und operative Gestaltungsimpulse für die Bauindustrie, München.

[9] Neuhäuser, T., Chen, Q., Rösch, M., Hohmann, A., Reinhart, G., 2020. Building Information Modeling im Fabriklebenszyklus. ZWF 115 (special), 66–69.

[10] 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, Berlin.

[11] Verein Deutscher Ingenieure, 2011. Fabrikplanung: Planungsvorgehen. Beuth 03.100.99.

[12] 2013. Verordnung über die Honorare für Architekten- und Ingenieurleistungen: HOAI.

[13] Jeffrey Ouellette. Software Implementations. https://technical.buildingsmart.org/resources/software-implementations/. Accessed 31 March 2021.

[14] Verein Deutscher Ingenieure, 2018. Building Information Modeling: Begriffe. Beuth 35.240.67.

[15] Patrick MacLeamy, 2004. Collaboration, Integrated Information, and the Project Lifecycle in Building Design and Construction and Operation. https://docplayer.net/61412208-Collaboration-integrated-information-and-the-project-lifecycle-in-building-design-construction-and-operation.html. Accessed 21 March 2021.

[16] Verein Deutscher Ingenieure und buildingSMART, 2020. Building Information Modeling: Informationsaustauschanforderungen. Beuth 35.240.67.

[17] 2019. The New Zealand BIM handbook: A guide to enabling BIM on building projects, Version 3.1 ed. [BIM Acceleration Committee], [Wellington], 1 online resource.

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 2552 Part 11.8 BIM – factory planning and spokesman of the buildingSMART Roundtable openBIM in factory planning.

Philipp Michaeli (*1982) received his doctorate in 2016 at the Technical University of Munich. From 2010 Dr.-Ing. Philipp Michaeli worked at ifp - Prof. Dr.-Ing. Joachim Milberg Institut für Produktion und Logistik GmbH & Co. KG and was employed there as managing director starting in 2015. He is also a member of VDI-guideline committee 2552 Part 11.8 BIM - factory planning.

Lisa Lenz (*1986) is a civil engineer and works as Post-Doc at 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 Management GLW GmbH and works as a freelance engineer.

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.

Fabian Matschinsky (*1988) is mep engineer and works as a project leader and BIM-coordinator at HL-Technik Engineering GmbH. His focus lies on the sustainability and digitalization of the building industry. He is leader of the BIM-department at HL-Technik and develops workflows to collaborate the various disciplines of Architecture, Engineering and Construction.

Franz Madl (*1959) is chief executive officer of pbb Architekten Ingenieure Ingolstadt and has more than 25 years of professional experience. He is founder of D.E.T. Bau - a group of experts with the aim of advancing digitization in the building industry and tries to establish BIM as well as digital design and construction as a standard in planning projects.

Rayk Henkelmann (*1977) is owner, manager and leading an engineering company for technical building equipment, specialized for electrical planning and consulting. He is member of the VDI and buildingSMART.

Conference On Production Systems And Logistics
CPSL 2021

2nd Conference on Production Systems and Logistics

Discovering Heuristics And Metaheuristics For Job Shop Scheduling From Scratch Via Deep Reinforcement Learning

Tilo van Ekeris[1], Richard Meyes[1], Tobias Meisen[1]
[1]*Chair of Technologies and Management of Digital Transformation, University of Wuppertal, Germany*

Abstract

Scheduling is the mathematical problem of allocating tasks to resources considering certain constraints. The goal is to achieve the best possible scheduling quality given a quality metric like *makespan*. Typical scheduling problems, including the classic *Job Shop Scheduling Problem* (JSP or JSSP), are *NP-hard*; meaning it is infeasible to use optimal solvers for big problem sizes. Instead, heuristics are frequently used to find suboptimal solutions in polynomial time, especially in real-world applications. Recently, *Deep Reinforcement Learning* (DRL) has also been applied to find solutions for planning problems like the JSP. In DRL, agents learn solution strategies for specific problem classes through the principle of trial and error. In this paper, we explore the connection between known heuristics and DRL: Heuristics always rely on features that can be extracted from the considered problem with low computational effort. We show that DRL agents, for which we limit the available observation to the underlying features of well-known heuristics, learn the behaviour of the more qualitative heuristics from scratch, while they do not learn the behaviour of less qualitative heuristics that would also be possible learning outcomes given the same feature as observation. Additionally, we motivate the use of DRL as a metaheuristic generator by training with the features of multiple basic heuristics. We show promising results that indicate that this learned metaheuristic finds better schedules in terms of *makespan* than any single simple heuristic – while only requiring simple computations in the time-critical solution phase and thus being faster than optimal solvers.

Keywords

Deep Reinforcement Learning (DRL); Production Planning; Scheduling; Job Shop Scheduling (JSP, JSSP); Proximal Policy Optimization (PPO); Heuristics; Metaheuristics

1. Introduction

In production and logistics, scheduling is the mathematical problem of allocating tasks to resources considering certain constraints. Different types of scheduling problems occur, the formulations range from static formulations like the *Job Shop Scheduling Problem* (JSP, JSSP) to formulations that include dynamics (e.g., newly appearing jobs [1] or machines with changing availability) or other complexities (e.g., the requirement to schedule tooling in addition to machines). Typical scheduling problems are combinatorial optimization problems and finding their solutions is proven to be *NP-hard* [2]. Classically, scheduling problems are solved by either optimal solvers, basic heuristics or metaheuristics. In this work we seek to apply *Deep Reinforcement Learning* (DRL) to scheduling problems. We focus on the following JSP formulation: The goal of the problem is to produce a schedule with minimal total production duration *(makespan)* by assigning a number of jobs n_j, divided into multiple tasks with fixed durations, to a fixed number of machines n_m (denoted a $n_j \times n_m$ JSP). The machine required for each task is given by the

DOI: https://doi.org/10.15488/11231 publish-Ing.

problem. The constraints for scheduling are: C1) the tasks within each job have to be scheduled in the given order; C2) only one task can be scheduled on each machine in each time-step. Figure 1a shows an example JSP with four jobs and four machines. An optimal solution regarding this JSP with *makespan* 24 is depicted in Figure 1b. Other optimal solutions can be trivially created by moving tasks which are not lying on the critical path (e.g., in Figure 1b, *Task 2-3* can be moved by up to three timesteps to the right without changing the resulting *makespan*).

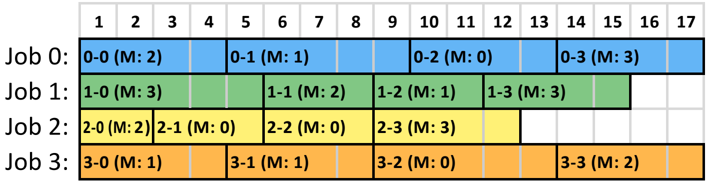


Figure 1a: Example of a 4 × 4 JSP: Each of the four jobs contains four tasks with a unique identifier ("id(j)-id(t)") with id(j) the job number and id(t) the task number within its job as well as the required machine ("M: id(m)") and the duration which is represented by the length in blocks

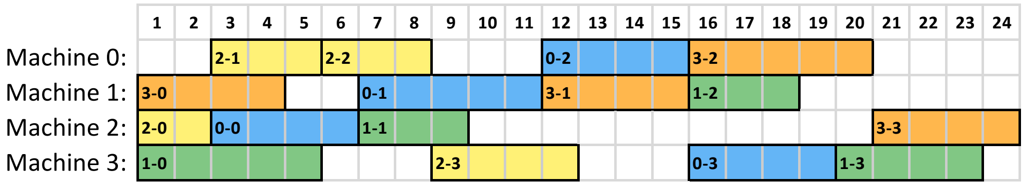


Figure 1b: (Example) optimal solution for the JSP from Figure 1a calculated with the Google OR tools solver

The objective of our study was twofold: First, we examined whether a DRL approach is able to discover basic heuristics for solving the JSP from scratch given a limited feature observation (Experiment E1). Second, we trained a DRL agent on a set of features to let the agent construct a scheduling strategy on these features. We then tested this strategy against basic heuristics and an optimal solver from a quality as well as from a run time perspective (Experiment E2).

2. Background and Related Work

2.1 Classic solution methods

For optimal solving, the JSP inputs and constraints can be expressed mathematically through a set of (in)equations and then put into state-of-the art solvers like the *CP-SAT solver* from the *Google OR Tools* suite [3], where CP indicates *Constraint Programming* and *SAT* the *Boolean satisfiability problem*. The *CP-SAT solver* uses a technique called *Lazy Clause Generation* [4] to find optimal solutions even to *NP-hard* problems as efficiently as possible.

Heuristics are frequently used, especially in real-world applications, where the available time for scheduling is often limited. Heuristics rely on features that are calculated from the problem with low computational effort so that heuristics' run times are typically low – but the simplifying nature of heuristics does not guarantee optimal or even good schedules in terms of a given quality metric. Research in heuristics is a vast field with a long history. Pinedo [5] discusses various approaches, from simple heuristics like *Earliest Due Date* (EDD), and *Earliest Release Date* (ERD) to other techniques such as *genetic algorithms* and *ant colony optimization*.

Finally, metaheuristics "combine basic heuristic methods in higher level frameworks aimed at efficiently and effectively exploring a search space" [6]. They try to get the best of two worlds: Use diverse features to try to use better planning decisions while maintaining a low computational effort.

2.2 Deep Reinforcement Learning

DRL is a machine learning discipline combining *Deep Learning* (DL) with *Reinforcement Learning* (RL) in which artificial agents are trained to take good actions given an observation based on the state of the environment [7]. Agents are trained by exposing them to a large amount of observations and giving them a reward signal after individual actions and/or a chain of actions (sparse rewards). The agents in DRL consist of at least one neural network that maps the observations to actions (cf. Figure 2).

DRL has recently seen the biggest breakthroughs in Games like classical board games Shogi, Chess and Go [8] as well as real-time strategy games like StarCraft II [9] because computer games and simulations allow to generate the large amounts of training data required by DRL more easily than real-world settings. Within the field of DRL, different training algorithms have been developed. The algorithm used in this work, *Proximal Policy Optimization* (PPO) [10], is widely used, for example by OpenAI in their work on the game DotA2, OpenAI Five [11]. PPO is a policy gradient algorithm which uses two separate neural networks, a value network and a policy network, as opposed to the algorithm *Deep Q-Network* (DQN), which uses only a value network [12].

Figure 2: Deep Reinforcement Learning approach

More recently, DRL has been applied to a diverse range of problems outside of games, including scheduling problems. Liu et al. [13] trained a DQN on the JSP and they formulated the agent's action as a choice from a set of fixed heuristics. This does not allow the agent to find its own heuristic from the underlying features. Luo [14] took a similar approach: DQN is used (with the enhancements *double DQN* and *soft target weight update*) and the agent's action consists of choosing from "six composite rules [which] are designed to simultaneously determine which operation to process next". Waschneck et al. [15] also apply DQN to production scheduling and construct a more complex multi-agent setup to solve their custom scheduling problem. The scheduling quality is not better than heuristics, but comparable to a human expert benchmark. Rinciog et al. [16] apply the AlphaGo Zero algorithm [17] to a JSP with multiple processing steps and compare the results to the simple heuristic *Earliest Due Date* (EDD) as well as to the classic search algorithm *Monte Carlo Tree Search* (MCTS). The results are better than EDD after Fine Tuning the RL agent by a small margin and better than MCTS by a larger margin (taking into account that MCTS was only run with a small number of roll-outs in the vast search tree of the scheduling problem). Finally, Kuhnle et al. [18] give a good overview about relevant research approaches and other related work. They also solve a dispatching problem with several groups of machines using the *Trust Region Policy Optimization* (TRPO) algorithm [19]. They extensively discuss the problem between sparse and dense rewards and describe the modelling of their RL agent in a detailed manner.

Although several groups of researchers have conducted experiments training DRL agents on scheduling problems, to our best knowledge there are no studies that explicitly examine what DRL agents learn given only basic feature observations with the connected heuristics.

3. Methods

This section describes the design of the DRL agent, the data generation, agent training and agent testing in the order that is typically followed when applying DRL to a specific problem.

3.1 Design of features and the agent's observation

In order to be able to learn, a DRL agent needs an observation that is calculated from the state of the environment (cf. Figure 2). In our experiments, this observation consists of a number (depending on the experiment) of feature vectors that are calculated from the environment state and each have one entry for each job, so the vectors have length n_j. In the following, we introduce the features used. Some are inspired by known basic heuristics and some are our own developments:

The *remaining job duration* RJD_j is defined by the sum of the durations of all unscheduled tasks from a job j (cf. Figure 3). Two well-known basic heuristics defined on this feature are: The *Longest Remaining Duration* (LRD) and the *Shortest Remaining Duration* (SRD), which at each iteration choose to schedule the next task from the job with the longest respectively shortest RJD. For example, the SRD heuristic would choose *Job 2* in Figure 3. Please note that a large number of heuristic rules can be defined on this feature (not only max/min).

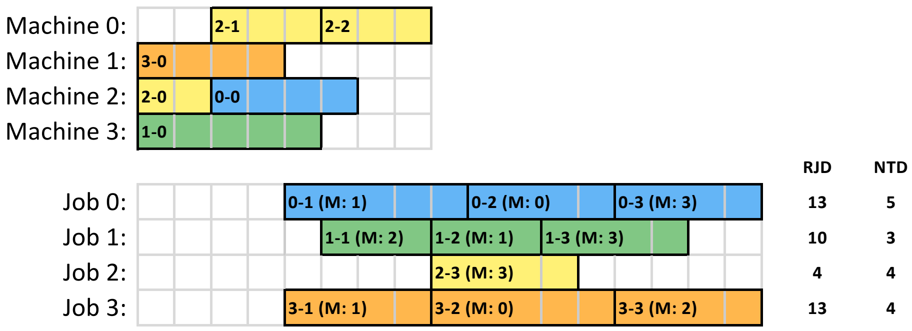


Figure 3: Partially scheduled example JSP from Figure 1a (top) with calculated features "remaining job duration (RJD)" and "next task's duration (NTD)" for the tasks remaining to be scheduled (bottom)

The *next task's duration* NTD_j is defined by the duration of the next task of a job j (cf. Figure 3). Two well-known basic heuristics defined on this feature are: The *Longest Processing Time* (LPT) and the *Shortest Processing Time* (SPT), which at each iteration choose to schedule the next task from the job with the longest respectively shortest NTD. For example, the SPT heuristic would choose *Job 1* in Figure 3.

The *remaining task count* RTC_j counts the number of unscheduled tasks of a job j.

The *bottleneck feature* BF_j seeks to guide the agent to choose tasks from jobs that mitigate machine bottlenecks. First, to determine a proxy of how likely a machine is to become a bottleneck, it sums up the duration of unscheduled tasks *per machine* (as opposed to *per job*). This vector is multiplied with a (n_m, n_j)-matrix where each item equals the sum of the unscheduled duration within job j to be scheduled on machine m.

3.2 Design of the agent's actions

In DRL, it needs to be defined how the agent is able to interact with the environment via actions (cf. Figure 2). We decided that the agent chooses from which job to schedule the next task. It cannot choose specific tasks; always the first unscheduled task from the chosen job is scheduled. Thus, the action is a number $a \in \mathbb{N}\,; 0 \leq a < n_j$. This action is then transferred to a scheduling component that schedules the resulting task at the earliest possible timestep complying with the two JSP constraints (cf. Section 1).

3.3 Design of the agent's reward

We seek to minimize the *achieved makespan* m_a of the completed schedule and the reward needs to be designed so that the agent is steered towards this goal. Because we can evaluate m_a only after the JSP is completely scheduled, we deal with so-called *sparse rewards.* The lowest possible or *optimal makespan* m_o of the problems is unknown (without running an optimal solver like the CP-SAT solver), so we compare m_a to a *lower bound makespan* m_{lb} which is calculated by simply adding all the task durations per machine and taking the maximum. Please note that $m_{lb} \leq m_o$ is satisfied for all possible JSPs, because $m_{lb} = m_o$ if and only if an optimal solution has no empty timesteps on the machine with the longest total duration and $m_{lb} < m_o$ otherwise. When m_a approaches m_{lb}, the reward shall rise, and with a greater slope for diminishing differences. We therefore decided to calculate the reward via a negative logarithm with the ratio between m_a and m_{lb} as argument. In order to leave room for punishing the agent for unwanted behaviour, we added constant shifts and scalings. The final definition of reward r is given in Formula (1) and plotted in Figure 4.

$$r = \left(-\ln\left(\frac{m_a}{m_{lb}} - 0.95\right) + 2.5\right) * 50 \qquad (1)$$

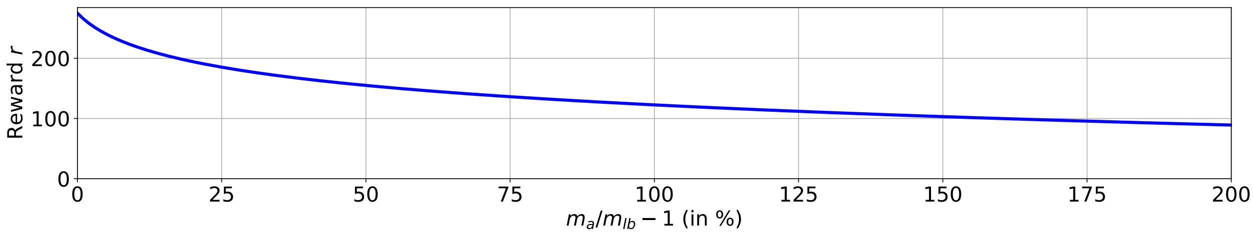


Figure 4: Reward r for successful solutions of the JSP as a function of m_a and m_{lb}

If the agent chooses a job that is already completely scheduled, it receives a large negative reward $r = -200.0$. The goal is that the agent learns that it should never choose such a job. This works in most of the cases, however, depending on the number of training timesteps, it is possible that the agent does not fully learn this behaviour which can result in an "action loop", where the agent always takes the "wrong" job while accumulating negative rewards. This behaviour was also described by Kuhnle et al. [18]. It can be mitigated by either detecting the action loop and forcing the agent to take a valid decision or choose an appropriate number of training timesteps for which this behaviour does not appear.

3.4 Problem classes and data generation

Depending on how exactly the JSPs are constructed, different problem classes are generated. As we know from extensive research in the ML community [20], the data that is used to train models and agents can (and usually will) introduce biases in the resulting behavior. We generate JSPs in a reproducible way, but we are aware that different choices could be taken that generate different JSP problem classes.

First, we initialize a random number generator with a seed. For a JSP with n_j jobs and n_m machines, we took the choice that the JSP is constructed so that each job also contains n_j tasks, so that the problems grow in two "dimensions" when augmenting n_j. For each task, the required machine is determined by sampling from a uniform distribution. For the duration d, we tested two different generation methods. In the first, we generate d by sampling from a normal distribution $d \sim \mathcal{N}(n_j, (0.25 \cdot n_j)^2)$ with mean n_j and standard deviation $0.25 \cdot n_j$ (discretizing and forcing $d \geq 1$). In the second, we sample d from a discrete uniform distribution $d \sim \mathcal{U}(1, n_j)$. This generates already two different problem classes of JSPs that we have both tested in our experiments.

3.5 Agent training

The training is done by exposing the agent to the generated training JSPs. In each timestep, the observation that is required for the experiment is calculated, normalized and passed to the agent. The agent chooses its action and might receive a reward directly (in the case of the negative reward due to choosing the wrong job), then receives the next observation. After the schedule for the current problem is finished, the agent receives the reward as described in Section 3.3 and is subsequently presented the next JSP. Our DLR agents are all trained with the PPO [10] algorithm. This algorithm trains two deep neural networks, a *policy network* and a *value network* (*Actor-Critic architecture*). For our problem sizes we kept the default network architecture of the Stable Baselines implementation [21], which defines the policy network and the value network as completely separate (no shared neurons) with two layers consisting of 64 neurons each. We kept the hyperparameters *learning_rate=0.0003*, *gamma=0.99* and *clip_range=0.2* at the proposed defaults. The agents were trained for 250,000 timesteps (one timestep being one *Reinforcement Learning* cycle of observation, action and reward) on JSP training data and tested on 1,000 JSPs from a different test set. Each JSP from the training set was only used once.

3.6 Baseline schedulers and agent testing

In order to test the DRL agent against baseline methods, we have implemented several other schedulers. To get the optimal solution, we use Google's OR tools JSP solver [3]. We also test against well-known standard heuristics (cf. Section 3.1) as well as other simple heuristics, like a random solver. Testing is done by letting all schedulers solve the JSPs from the test set and comparing the resulting schedules. To compare the scheduler similarity, we compare the produced schedules one-by-one and calculate their distances by taking the sum of absolute differences of the task starting times (in discrete timesteps) for each task, divided by the total number of tasks. We call this distance measure *cumulative absolute task deviation* (CATD). To examine the absolute and relative scheduling quality, we measure the *achieved makespan* m_a for every schedule and for each scheduler calculate the mean, minimum, maximum and standard deviation of the distribution of *makespans* m_a and run times (excluding set-up times for all algorithms).

4. Experiments and Results

4.1 Discovering basic heuristics from basic features from scratch (Experiment E1)

In this experiment we examine if the DRL agent is able to discover known basic heuristics from scratch when given only the heuristics' feature as observation and which of the possible heuristics it learns. We have run the experiment (independently) for the two features RJD and NTD (cf. Section 3.1). Figures 5 and 6 show the distances of schedules (cf. Section 3.6) for a selection of scheduler combinations.

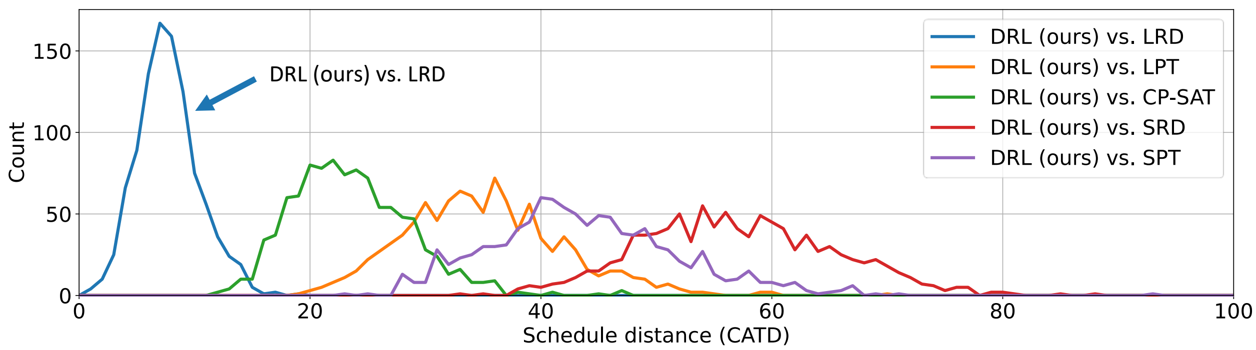


Figure 5: Schedule distances (CATD) histogram comparing our DRL scheduler (trained with the RJD feature as the only observation) against baseline schedulers

Each graph is a histogram that shows the distances between the schedules of two schedulers. A graph that is located more towards the origin means that more schedules generated by the two schedulers for the same problem are similar, while a graph located more towards larger arguments means that the schedules are less similar. The results indicate that the DRL agent is able to discover known heuristics from scratch given the underlying basic feature. For the feature RJD, the agent learns the heuristic LRD and not SRD, which both rely on the same feature, presumably because LRD is the better performing heuristic in terms of *makespan* (cf. Section 4.2). For the feature NTD, the similarity of the DRL scheduler and the LPT scheduler is greater than for any other heuristic (including SPT), but the results are less obvious than in the RJD case. Our explanation for this is that for the feature NTD, neither the heuristic LPT nor the heuristic SPT produce really good schedules (in terms of *makespan*, cf. Section 4.2) and the DRL agent thus is not rewarded to either clearly learn the behaviour of LPT nor SPT.

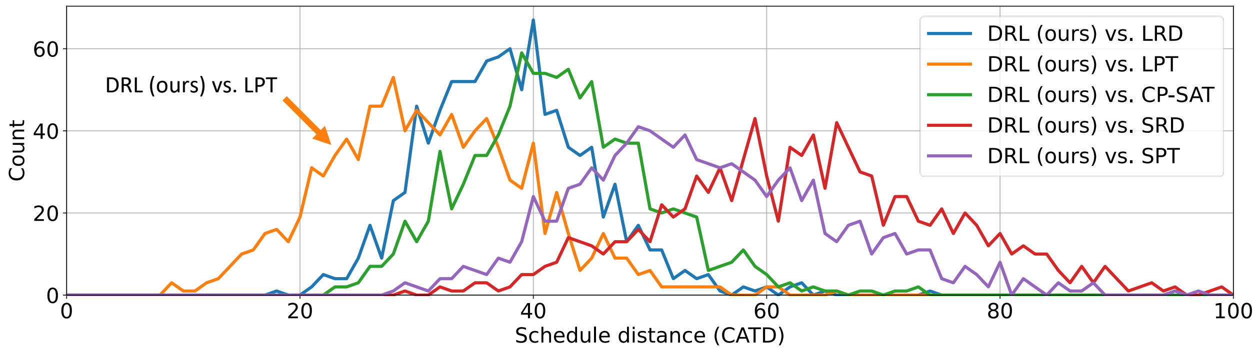


Figure 6: Schedule distances (CATD) histogram comparing our DRL scheduler (trained with the NTD feature as the only observation) against baseline schedulers

4.2 Learning a new metaheuristic from multiple basic features from scratch (Experiment E2)

In this experiment we trained DRL agents on JSP problems of sizes 8x8 and 12x12 using a set of features as the agent's observation (as opposed to experiment E1, where the agent's observation only consists of a single feature in each experiment). An exhaustive study has been performed in which we have found that an observation consisting of the three features RJD, RTC and BF leads to the best results (in terms of *makespan*). See Tables 1 and 2 for the results on 8x8 and 12x12 JSPs respectively, both with normal task duration distribution (cf. Section 3.4).

Table 1: *Achieved makespans* m_a and solver run times on 1000 8x8 JSPs

	Achieved makespan m_a				Run time per JSP (in ms)			
	mean	min	max	stdev	mean	min	max	stdev
CP-SAT (optimal)	107.9	86	164	10.9	64.3	27.6	412.3	28.8
DRL (ours)	**120.9**	**91**	**169**	**10.9**	**298.5**	**257.3**	**823.0**	**32.6**
LRD	121.3	96	168	10.9	11.0	9.7	22.3	1.0
Random	137.3	108	230	13.2	2.2	1.9	5.8	0.2
LPT	142.6	107	198	14.3	3.0	2.6	5.2	0.3
SPT	148.9	109	231	16.5	8.6	7.7	15.7	0.7
SRD	158.2	115	226	17.8	17.0	15.1	183.8	6.7

The results indicate that learning a good metaheuristic from scratch from a set of basic features is possible. From a quality perspective, we achieve *makespans*, that are on average lower (better) than those from the best single heuristic. From a run time perspective, we achieve lower run times than the optimal solver for 12x12 JSPs. When growing the problem size from 8x8 to 12x12, we see more than 20x increase in the mean run time of the optimal solver, but less than 3x increase for the DRL solver. We argue that this difference is due to the *NP-hardness* of the problem and that the effect would amplify with growing problem sizes. We thus find a compromise position of solution quality and run time that lies between heuristics and optimal solver.

Table 2: *Achieved makespans* m_a and solver run times on 1000 12x12 JSPs

	Achieved makespan m_a				Run time per JSP (in ms)			
	mean	min	max	stdev	mean	min	max	stdev
CP-SAT (optimal)	236.8	200	371	17.5	1518.3	123.2	33825.7	2608.4
DRL (ours)	**275.4**	**233**	**376**	**18.1**	**763.6**	**712.1**	**1981.2**	**58.8**
LRD	277.2	234	391	18.5	39.0	36.2	76.9	2.9
Random	316.0	258	419	23.9	5.8	5.1	10.4	0.4
LPT	337.3	261	455	28.6	7.6	6.7	11.8	0.5
SPT	350.3	276	519	30.1	17.5	16.0	27.9	1.2
SRD	374.0	294	540	32.6	48.9	45.4	82.2	3.3

Please note, that the run times between the simple heuristics LRD and SRD (analogously LPT and SPT) result from slightly more complex code to obtain a minimum value while ignoring values that equal zero in comparison with obtaining a maximum. We found that just putting more of the simple features to the DRL agent does not help to find schedules with a better quality (the contrary is true). Intuition suggests, that the DRL agent is capable to completely learn which features to weigh more or less so that more features should not decrease the scheduling quality. But at least this is not the case if the number of training timesteps and/or the network architecture/size remain unchanged. We did not observe significant differences between the two tested problem classes (cf. Section 3.4), even if an agent was trained on data with normal duration distribution and tested on data with uniform duration distribution.

5. Discussion and Outlook

In this paper we have shown two distinct results: First, a DRL agent is able to learn basic heuristics from scratch given underlying basic features. The results indicate that the agent behaves similarly to the best performing heuristic that can be defined on the feature. Second, a DRL agent can discover a good metaheuristic automatically, given multiple basic features. It generates lower *makespans* than any single simple heuristic. While this solution method is not competitive compared to heuristics from a run time perspective, it requires much less computational effort than an optimal solver. Of course, we would like to improve the quality margin of the DRL agent compared to the best simple heuristic. This is important to be able to completely justify the use of DRL with its higher runtime and additional training effort for scheduling problems. Thus, one path to follow is to increase the number of basic features that are presented to the DRL agent, trying to find the features that lead to better results. We would like to examine the JSPs that can currently only be solved with large errors in comparison to the optimal solver in order to learn something about the structure and to produce new feature vectors that might help the agent. Another idea is to use generative processes (e.g., *Generative Adversarial Networks*, GANs) to produce more of those "hard" JSP instances so that we can augment the amount of those JSPs in the training data set with the goal that the DRL agent can learn more from those instances. We would also like to further evaluate the impact of other reward function definitions.

The decision logic of the metaheuristic discovered by the DRL agent is hidden inside a neural network and could be made visible via "explainable AI" techniques in future work.

Currently, we must apply feature engineering to select the best set of features as the agent's observation. We would like to approach *end-2-end DRL* for scheduling problems, meaning to find a representation that describes the structure of the problem directly and can be fed into the DRL agent.

In our approach, we can possibly generalize to the number of machines, but not to the number of jobs, because the agent requires the same observation vector size as well as a fixed action vector size (both depending on the number of jobs) determined during training. In the future we would like to look into applying *Graph Neural Networks* (GNNs, like in [22] or [23] for the TSP) to solve JSPs, because the GNN architecture offers the possibility to generalize to larger problem instances.

Our far-stretched goal is to apply DRL to real-world production scheduling scenarios. Therefore, we would also like to approach dynamic scheduling problems (like in [1]) in the future, in which new jobs arrive or jobs can get deleted dynamically and machines can have defects so that they cannot be used in certain timesteps. Currently, we only consider a reward solely based on *makespan*. But in real-world scenarios, other goals come into play. Specifically, in the future we would like to include plan robustness into the reward. Our agent should not only be able to minimize *makespan*, but also produce schedules that are subjected to minimal change in the case that jobs, machines or other resources become unavailable. Finally, we would also like to tackle more complex scheduling problems that exist in the real world, e.g., including tooling changes within production that need to be scheduled along with the job scheduling.

Acknowledgements

This research work was undertaken within the research project AlphaMES [24] funded by the German Federal Ministry for Economics and Energy (BMWi).

Appendix

The Python framework *jobshop* that includes all our design choices in code and can be used to produce all the described results and diagrams will be published under an open source license upon publication.

References

[1] Fang, J., Xi, Y., 1997, A rolling horizon job shop rescheduling strategy in the dynamic environment, The International Journal of Advanced Manufacturing Technology, 13/3:227–232, DOI:10.1007/BF01305874.

[2] Gonzalez, T., 1982, Unit Execution Time Shop Problems, Mathematics of Operations Research, 7/1:57–66, DOI:10.1287/moor.7.1.57.

[3] Perron, L., Furnon, V., 2019, OR-Tools Ver. 7.2. Google, [Online]. https://developers.google.com/optimization/.

[4] Stuckey, P. J., 2010, Lazy Clause Generation: Combining the Power of SAT and CP (and MIP?) Solving, in Integration of AI and OR Techniques in CP for CO Problems, Berlin, Heidelberg, pp. 5–9.

[5] Pinedo, M. L., 2016, Scheduling: Theory, Algorithms, and Systems, 5th ed. Springer International Publishing.

[6] Blum, C., Roli, A., 2001, Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison, ACM Comput. Surv., 35:268–308, DOI:10.1145/937503.937505.

[7] Sutton, R. S., Barto, A. G., 2018, Reinforcement Learning: An Introduction. Cambridge, MA, USA.

[8] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., et al., 2018, A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, Science, 362/6419:1140–1144, DOI:10.1126/science.aar6404.

[9] Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., et al., 2019, Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature, 575/7782:350–354, DOI:10.1038/s41586-019-1724-z.

[10] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., Klimov, O., 2017, Proximal Policy Optimization Algorithms, arXiv:1707.06347 [cs].

[11] OpenAI, Berner, C., Brockman, G., Chan, B., Cheung, V., et al., 2019, Dota 2 with Large Scale Deep Reinforcement Learning, arXiv:1912.06680 [cs, stat].

[12] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., et al., 2015, Human-level control through deep reinforcement learning, Nature, DOI:10.1038/nature14236.

[13] Liu, C.-L., Chang, C.-C., Tseng, C.-J., 2020, Actor-Critic Deep Reinforcement Learning for Solving Job Shop Scheduling Problems, IEEE Access, 8:71752–71762, DOI:10.1109/ACCESS.2020.2987820.

[14] Luo, S., 2020, Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning, Applied Soft Computing, 91:106208, DOI:10.1016/j.asoc.2020.106208.

[15] Waschneck, B., Reichstaller, A., Belzner, L., Altenmüller, T., Bauernhansl, T., et al., 2018, Optimization of global production scheduling with deep reinforcement learning, Procedia CIRP, 72:1264–1269, DOI:10.1016/j.procir.2018.03.212.

[16] Rinciog, A., Mieth, C., Scheikl, P. M., Meyer, A., 2020, Sheet-Metal Production Scheduling Using AlphaGo Zero, DOI:10.15488/9676.

[17] Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., et al., 2017, Mastering the game of Go without human knowledge, Nature, 550/7676:354–359, DOI:10.1038/nature24270.

[18] Kuhnle, A., Kaiser, J.-P., Theiß, F., Stricker, N., Lanza, G., 2020, Designing an adaptive production control system using reinforcement learning, Journal of Intelligent Manufacturing, DOI:10.1007/s10845-020-01612-y.

[19] Schulman, J., Levine, S., Moritz, P., Jordan, M. I., Abbeel, P., 2017, Trust Region Policy Optimization, arXiv:1502.05477 [cs].

[20] Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A., 2019, A Survey on Bias and Fairness in Machine Learning, arXiv:1908.09635 [cs].

[21] Raffin, A., Hill, A., Ernestus, M., Gleave, A., Kanervisto, A., et al., 2019, Stable Baselines3. https://github.com/DLR-RM/stable-baselines3.

[22] Zhang, C., Song, W., Cao, Z., Zhang, J., Tan, P. S., et al., Learning to Dispatch for Job Shop Scheduling via Deep Reinforcement Learning, p. 12.

[23] Joshi, C. K., Cappart, Q., Rousseau, L.-M., Laurent, T., Bresson, X., 2020, Learning TSP Requires Rethinking Generalization, arXiv:2006.07054 [cs, stat].

[24] van Ekeris, T., AlphaMES - TMDT – Bergische Universität Wuppertal. [Online]. https://www.tmdt.uni-wuppertal.de/de/projekte/alphames.html.

Biography

Tilo van Ekeris (*1986) has graduated with a Diploma degree from RWTH Aachen in Computer Engineering and a M. Sc. degree from Ecole Centrale Paris. He worked for 8 years for a top management consultancy company and several IT companies. Tilo worked at the TMDT during 2020/21 where he was involved in a range of Machine Learning/Artificial Intelligence projects.

Richard Meyes (*1989) has graduated as M.Sc. in Physics from RWTH Aachen and has worked as a research scientist in neuroscience and medicine as well as artificial intelligence and deep learning applications for industrial use cases for more than seven years. He is currently research group leader at the TMDT for explainable AI for industrial sensor data analytics.

Tobias Meisen (*1981) is Professor of Digital Transformation Technologies and Management at the University of Wuppertal since September 2018. He is also the Institute Director of the In-Institute for Systems Research in Information, Communication and Media Technology, the vice-chair of the Interdisciplinary Centre for Data Analytics and Machine Learning and a co-founder of Hotsprings GmbH.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Pricing For Smart-Product-Service-Systems In Subscription Business Models For Production Industries

Tobias Leiting[1], Günther Schuh[1], Volker Stich[1], Jana Frank[1]
[1]Institute for Industrial Management (FIR) at RWTH Aachen University; Germany

Abstract

In the production industry, subscription business models have the potential to create long-term relationships where a supplier provides a continuous value-oriented service to a customer based on digitalisation. Monetising this increase in value through pricing represents a central challenge for suppliers in subscription business. Unlike the current dominant transactional business, the focus of pricing is on the value-in-use of the customer (e.g. on the increase in output for the customer). In this regard, there is so far no pricing approach for practice that allows the linking of the performance data of the customer with the periodically charged price. However, in subscription businesses, such an approach is required to create win-win situations for the customer and supplier through continuous performance improvement. Therefore, this paper develops a novel process model for pricing of smart-product-service-systems in subscription business for production industries. This process can serve as basis for suppliers of subscriptions in the production industry to align pricing with the created value-in-use. In the long term, this allows companies to systematically develop their pricing to monetise the potential of digitalisation.

Keywords

Digitalisation; Subscription business; Value-based pricing, Value-in-use; Data-based pricing; Smart-product-service-systems

1. Motivation and challenge

In the production industry, smart services are increasingly being offered alongside physical products and services. These digital offerings provide the potential to raise flexibility, efficiency, productivity, and quality within producing companies to a new level. However, despite great expectations and investments, digitalisation has not led to a significant increase in sales for machine and plant engineering (suppliers) and value creation for manufacturing companies (customers) in the last decade [1]. One of the reasons for this is that investment services and goods in the production industry are provided by suppliers mainly in transactional business until now [2]. In transactional business, suppliers focus on offering products and services in single transactions and leave the responsibility for the operation during the usage phase to the customer. However, digital solutions are designed to provide a continuous solution during the usage phase. The customer benefits particularly when the supplier integrates into their processes and continuously provides resources during the operation in the usage phase [3]. To achieve commercial success with the digitalisation of the production industry, tailored business models must be implemented. For this purpose, first industrial companies transformed their previous transaction-based business models into subscription business models in recent years. Prognoses indicate average annual growth rates of 23 percent for suppliers in this sector [4]. In these subscription business models, the supplier and the customer establish long-term relationships through a subscription contract. The focus of the supplier is no longer on selling isolated products or services transactionally, but on permanently satisfying customer needs through cooperation and collaborative improvement towards a solution tailored precisely to customer requirements [5]. For this purpose, a smart-product-service-system (solution) is offered in the form of a customer solution that integrates a connected product, services and smart services into a single offering for the customer, which is continuously adapted to the customer and improved over the duration of the usage phase [1]. The overall objective is to create a win-win situation in which a customer realises substantially higher added value through the smart-product-service-system and the supplier generates substantially higher revenues and profits from the customer over the usage phase by continuous subscription pricing [4].

The central task of pricing is thereby to implement functions and mechanisms that allow a supplier to participate in the provided added value for the customer during the usage phase [6]. Herby, the suppliers are faced with the challenge of pricing individual, continuously adapted solutions accurately as part of novel subscription business models that focus on the usage phase of the customer. If the pricing is not configured in the favour of all parties or is not linked to the value provided to the customer, the customers will not accept it and the suppliers have no incentive to continuously improve the provided value. As a result, suppliers do not invest resources in improving performance. Consequently, potential for increasing customer value is not exploited and suppliers do not participate in customer value, which is frequently many times higher than the effort required to improve performance [6,7,5,8]. Digitalisation acts as a key enabler for individual value recording due to access to usage data. A scalable implementation of value-based pricing can be facilitated by innovative pricing models where the price is calculated based on the usage data from the customer [9]. To create these mechanisms systematically and efficiently, suppliers need a structured approach as well as defined tools and a decision-making basis. The cost-plus or competition-based pricing approaches which are established in industry for transactional product sales are not applicable for subscription business, as they do not consider the value of the customer in the usage phase [4,9]. Furthermore, existing subscription pricing approaches from other sectors cannot be adopted to the manufacturing industry due to the complexity and high degree of interaction and individualisation within the decision-making processes, smart-product-service-systems and business relationships [10]. Accordingly, the aim of this paper is to develop such an approach for pricing of smart-product-service-systems for subscription business in the production industry. To ensure the scalability of this approach in practice, it is designed as an operation-oriented process model that integrates data and information from the usage phase of the customer.

2. Research ambition

From a research perspective, implementing such a model for pricing of smart-product-service-systems in the production industry leads to a novel value-based pricing approach, which has been not possible in the transaction-based business so far. In existing value-based pricing for transactional business, all activities take place during the sales phase and the pricing is focused on a specific point in time before the actual use of the solution [11,12,9]. The determination of the value for the pricing is done a priori regarding the usage phase. Consequently, the pricing is based on an anticipated value for the customer (s. figure 1). This means, that the generated value-in-use has no direct impact on the price paid, so that there are no direct monetary consequences for the supplier from the usage phase. Accordingly, no direct incentives are created for the supplier to invest resources in the continuous improvement of the solution for the customer and thus in increasing the customers value. In contrast, the subscription business model is focused on continuously recurring payments during the usage phase [5]. The pricing is applied in the usage phase for a continuously provided solution, based on a contractual framework defined in the sales phase. As result, the pricing takes place partially in the usage phase. This means that subscription pricing between a supplier and a customer in production industries can be divided into two central phases with different terms and conditions. First phase is the sales phase in which a subscription contract is worked out between the supplier and the customer. This contract must specify the underlying pricing model, the pricing components, and the pricing metrics

[13]. After conclusion of the contract, the second phase is the usage phase of the customer. Within this phase, the solution provided in a previously defined cycle and the resulting value-in-use are priced on a recurring basis. Therefore, the prices are calculated based on the recorded data of the customer according to the conditions specified in the contract. Through this mechanism, it is possible to realise an a posteriori pricing based on the value-in-use at the customer. On the one hand, pricing must create incentives for continuous improvement for the supplier and on the other hand, the dynamically changing value-in-use of the continuously adapted solution must be recorded based on customer data. Therefore, pricing in subscription models within the usage phase is a flexible and iterative process [4,5].

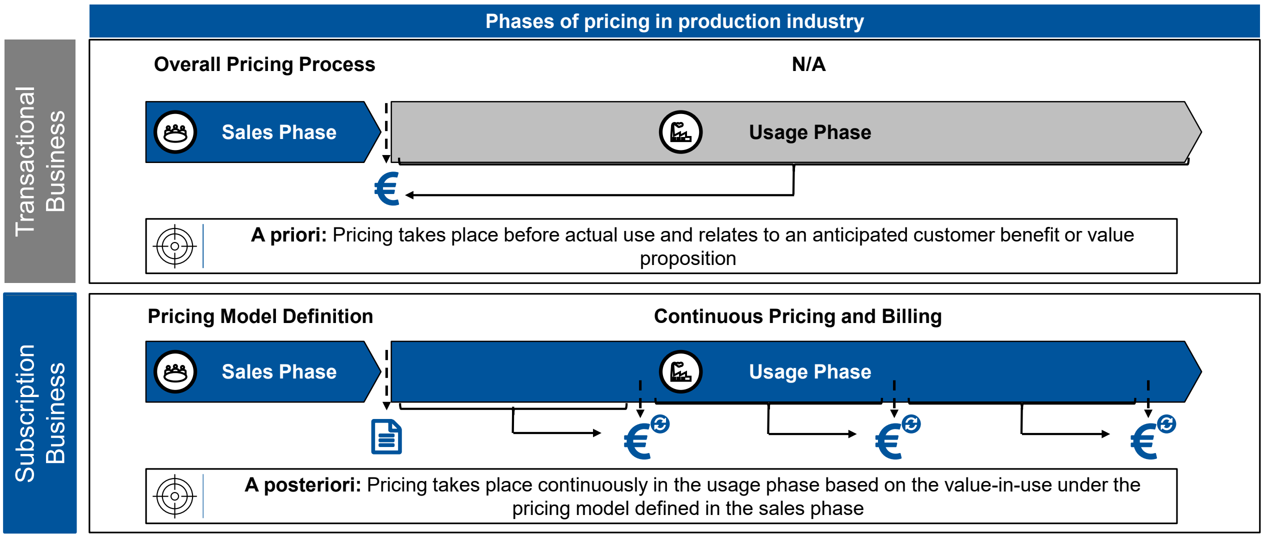


Figure 1: Phases of pricing in production industry (own figure)

Smart-product-service-systems in subscription business represent a complex integrated customer solution consisting of product, service, and smart services, which are individually tailored to a customer [1]. A novel aspect of subscription business compared to existing value-based pricing from the production industries [11,12], is that these represent individual solutions from the perspective of the customer. This involves taking on the technical and economic risks from the customer. In this context, the customer is offered the availability, the usage, the result, or the success with the offered solution [10]. In order to perform pricing, it is necessary to determine how the provided solutions are individually perceived by the customer [14]. A central feature in this regard is the measurement of the value-in-use. This value, which is mainly referred to conceptual approaches for the evaluation of value propositions of services, is not used in operational approaches for pricing [15,16]. In practice, companies have difficulty in identifying and quantifying the key value drivers for individual customers. Moreover, value-in-use is inherently phenomenological, meaning that at one customer, different stakeholders evaluate the value of the same solution differently [15]. This implies that the value assessment for pricing must be based on defined and measurable data [16]. Additionally, for subscription business, the selection of the pricing model and the pricing metric is significantly more complex than in the transactional business. The conventional price metric in transactional business reflects the logic of exchanging a product for a fixed one-time price [9]. To capture customer and supplier requirements, the price metrics in the subscription business are usually multi-dimensional and a price is composed of fixed elements as well as variable elements [13]. In particular, the variable component could be linked to the operational data of the customer. A supplier has a high degree of freedom in determining the price dimension and price metrics, by means the design represents a complex decision-making situation [9,8]. For this purpose, the way of data recording and including in the pricing must be specifically defined. Until now, existing approaches do not link price metrics with specific data from industrial operations of a customer. In summary, to achieve the objective of this paper, the comprehensive extension of current value-based pricing approaches is necessary. Therefore, the process model to be developed goes beyond the existing research by addressing both the sales and the usage phase for subscription business models. In this context, novel functions and activities for pricing of smart-product-service-systems in subscription business for production industries must be defined.

3. Research method and framework

For the development of the new process model, the research method of modelling is used. The method follows the approach of systems engineering and the principles of orderly model design [17]. For this purpose, the method is based on five key principles. 1. Principle of accuracy: The model must have semantic and syntactic correctness. 2. Principle of relevance: Only the essential aspects that provide a benefit for the task should be included in the model. 3. Principle of clarity: The model must be understandable and intuitive for a user from the practice. 4. Principle of comparability: The recommendations of the process model for pricing must also be usable independently of the application of the model. 5. Principle of systematic structure: The model should have defined interfaces to corresponding models.

For designing a process model for the operative execution of business processes in compliance with these key principles, frameworks are used in research practice. Frameworks define connections among individual elements of the methodology on a high level of abstraction and arranges them in a hierarchy or sequence [17]. To develop such a framework, relevant elements from processes of existing pricing approaches provide the basis. A structured literature research was carried out to identify relevant existing process models. As result, approaches for pricing of subscription were identified. Most of these approaches [18,13,19,6] are primarily designed for pricing in subscription business in the software industry. Therefore, approaches for value-based pricing from the transaction business of the production industry are included as well [20–22,11]. The analysis of existing approaches revealed a consistent pattern regarding the content and process of pricing (s. figure 2). Although the approaches differ to some extent in their scope and design, the elements of all models can be clearly assigned to a higher-level process framework. This process framework contains of the four superordinate elements reference basis, value assessment, price dimension and price metric.

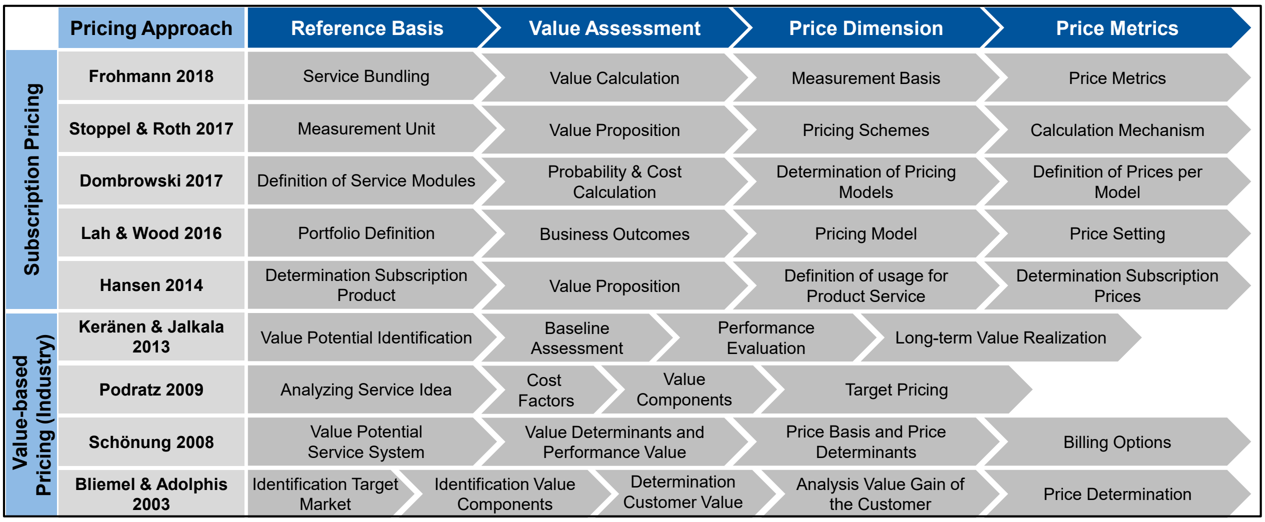


Figure 2: Definition of the framework for the process model to be developed (own figure)

The first element is the **reference basis** in which the offered solution is recorded, defined, and described. This element is particularly important for complex solution offerings in the production industry [11]. In **value assessment**, a recording of the value of the solution from the perspective of the customer and the supplier is conducted. For the customer, this is defined by the value-in-use [9,23] and for the supplier by the customer lifetime value [24].The **price dimension** describes a qualitative logic between the value of a solution and the charged price for the customer [13,10]. The definition of price elements is conducted, that relate a monetary equivalent to an offered solution component that is accepted by both parties [9]. The **price**

metric establishes a quantitative relationship between the charged price and the data-based performance parameters from the customer [13]. Through a specific calculation mechanism, the charged price for the solution is quantified with the data.

4. Development of a pricing approach for smart-product-service-systems

In this chapter, the process model for pricing in subscription business for production industries is derived. To this end, chapter 4.1 develops the eight process steps of the process model based on the subscription and pricing characteristics. Subsequently, chapter 4.2 characterises the individual process steps with features and properties.

4.1 Process model for pricing in subscription business for production industry

The process model to be developed is formed by the previously developed framework for pricing with the four elements reference basis, value assessment, price dimension, and price metric as well as the two phases of pricing for subscription business, the sales phase and the usage phase. The two phases are divided by the signing of the subscription contract. The steps of the usage phase are repeated over and over when the subscription is renewed. As the subscription model in many cases is constantly modified, the steps of the usage phase are continuously processed and flexible [4,5]. During both phases, each element of the pricing framework is relevant and addressed by requirements and tasks. The process model contains an objective-based business process with a detailed workflow logic consisting of eight process steps a supplier in practice performs to determine the specific price for a smart-product-service-system (s. figure 3).

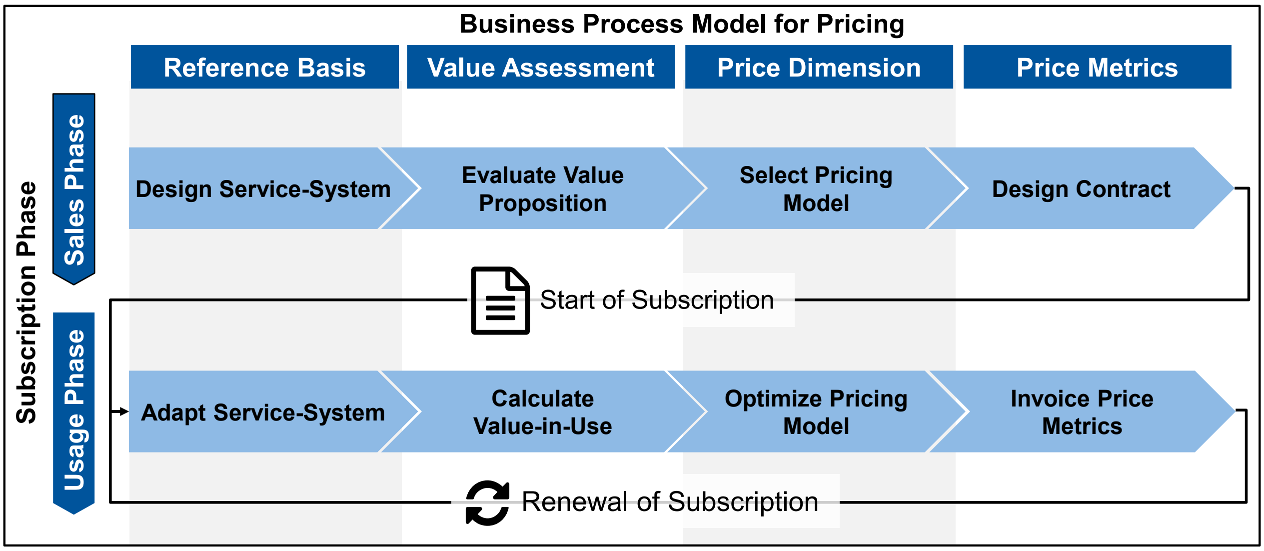


Figure 3: Process model for pricing in subscription business (own figure)

The objectives within each phase result from the elements of the pricing framework, the information and data available at the phase, and the characteristic features of subscription business models [2]. The characteristic feature of reference basis is the continuous increase in value through performance enhancement. The offered solution is oriented to the value-in-use for the customer. The supplier is focused on integrating the solutions into the value creation processes of the customer. A solution is provided individually for a customer and is based on the individual requirements of a customer. The sales phase begins with the **design** of individual **service-systems** as a solution that is tailored to the customers' needs. During the usage phase, the solution is continuously optimised on a higher customer value by the continuous **adaption** of the **service-systems**. In value assessment the characteristic feature is the knowledge of individual customer value and needs. To address individual value propositions, an **evaluation of the value proposition** is conducted in the sales phase. A digital access to the customer in the usage phase enables recording of individual usage data. This provides the supplier an in-depth knowledge of the of the customer for the **calculation of the value-in-use** of the customer. Characteristic feature of the price dimension is an integrated, continuous, and participatory value creation. A long-term, participative co-creation is established between the customer and the supplier. The basis for this is established by **selecting the right pricing model** in the sales phase. In order that the customer profits by improving the performance processes and the supplier profits through an increase in revenue of the customer, an **optimisation of the pricing model** needs to be conducted within the usage phase for the exploitation of bilateral benefits. Characteristic features for the price metrics are periodic payments based on generated value. Within the sales phase a **contract is designed** that specifies the price components and the reference basis for the chosen pricing model. The continuous transfer of the solution from the supplier to the customer takes place in the usage phase. In return, the customer pays periodic fees to the supplier that are linked to the provided value. For **invoicing the price metrics,** the data of the defined price reference basis are recorded, and the price is calculated and charged for each payment period.

4.2 Properties of the process model for pricing in subscription business

Based on the defined process model, properties for each step are derived and defined in the following. These properties structure the tasks within each phase. An overview of the properties within each step is shown in the following figure (s. figure 4).

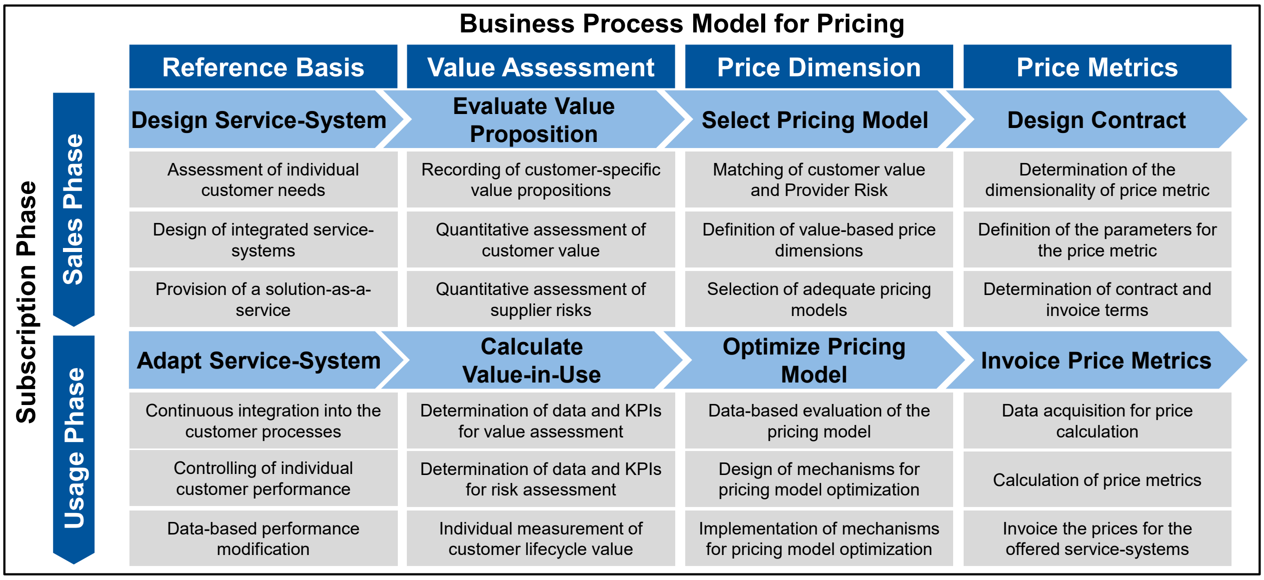


Figure 4: Properties of the process model for pricing in subscription business (own figure)

1. Design Service-Systems: The process of pricing starts with the customer and their needs. The supplier must perform an assessment of the individual customer needs. For this purpose, a supplier must analyse and understand the objectives, processes, resources, and activities of the customer [25]. The second sub-step is to transform the individual customer needs into a concrete solution specification [25]. Based on the specifications, products, services, and smart services must be generated for the design of integrated service-systems for the customer. For the successful design of product-service-systems, Belz defines six design principles. These are the integration principle, the principle of accounting, the participation and explanation principle, the evolution principle, the long-term principle, and the relevance principle [14,12]. Products, services, and smart services are subsumed under a superordinate customer solution, which is integrated into the customers value chain in relational processes [25]. The service-system is provided to the customer as a solution-as-a-service that meets his requirements. Products act as a distribution channel for the provided solution [15].

2. Evaluate Value Proposition: According to the service-dominant logic, a supplier alone cannot offer values to a customer, but only value propositions. Value for the customer arises from the co-creation of value in which both the customer and the supplier are involved [15]. To record the customer-specific value propositions, it is necessary to identify which preferences and objectives the customer has in the dimensions of efficiency, cost, and quality [3]. In the next sub-step, relevant customer data for the quantitative assessment of this value propositions must be identified and recorded. This data is used to determine a reference point before the subscription solution is provided to the customer [21]. Based on usage data from other customers of a comparison group, a quantitative assessment of the expected customer value income must be carried out [3]. Smart-product-service-systems in the production industry that include investment products involve associated risks for the supplier in the form of counterparty default, market price, and operational risks. Analogously to customer value, a quantitative assessment of the supplier risk must be carried out based on available data [1].

3. Select Pricing Model: The choice of the pricing model has a major impact on customer and supplier value realisation. To select the price dimension, an evaluation must be carried out, where a matching of the customer's value and the supplier's risk is made [23]. On the one hand, taking on promises of higher value leads to more customer value and on the other hand, taking on more risks of the customer can lead to high costs, if the supplier cannot influence them [10]. For the offering in subscription models, a distinction can be made between four different solution types. These successive types define the offer of availability, usage, result, or economic success for the customer. The definition of the value-based price dimension is conducted by selecting the offered solution type. The selection of adequate pricing models is then carried out in accordance to this. For the availability type, the price is taken for the available time of the solution (e.g. flat rate). In the usage type, a price is linked to used time units (e.g. pay-per-use). In the result type, a price is based on the produced unit (e.g. pay-per-output) and in the success type, a price is based on the increase in performance (e.g. profit sharing) [10].

4. Design Contract: For the continuous payments in subscription models, there are different options for the supplier in terms of pricing components. For the determination of the dimensionality of the price metric, a differentiation can be made between one-dimensional metrics with a variable or fixed component, two-dimensional metrics with a variable and a fixed component, and three-dimensional metrics with an additional component at the beginning of the subscription [13]. The selected price models need to be designed using price metrics based on customer-specific data [9]. Criteria for the definition of the parameters for pricing metrics are goal orientation, stakeholder-driven, and simple measurability [9]. The sales phase completes with the determination of contract and invoice terms. For the invoice terms, a distinction can be made between ex-ante billing and synchronized ex-post billing. Synchronized billing is particularly suitable for pricing based on the usage data from a customer [13,11].

5. Adapt Service-Systems: The supplier integrates the solution forward into the value creation process of the customer to link the processes of the two stakeholders [10]. A continuous analysis of individual customer needs is required. The supplier aligns the services, resources, and processes to these needs. Additional solutions are configured under continuous supplier integration into the customer processes. Both the supplier and the customer provide data, resources, and activities as co-producers [15]. Criteria of successful supplier integration are the focus on benefit processes, benefit development at the point of use, the integration paradigm, the interaction paradigm, and the combined push-pull principle [15]. The aim is to offer optimised solutions to the customer in close cycles of performance optimisation. To achieve this, conclusions are drawn about the actual usage behaviour of the customer within an iterative, data-driven learning process. Based on that customer insights, a data-based performance modification is carried within continuous releases [3].

6. Calculate Value-in-Use: Using KPIs, strategic success factors of a company can be quantified. KPIs serve for operationalisation, specification, stimulation, management, and control functions. For pricing, these indicators are primarily used to operationalise and specify the value-in-use [6]. To do this, KPIs and data must be determined for the quantitative value assessment of the customer. The data for this mostly comes from the systems and processes of the customers and from parameters that are defined specifically for the calculation of the KPIs. As a further factor, the profitability of a business must also be ensured by taking risks into account [9]. Here, suitable data and KPIs must be determined for risk assessment as well. In this context, data for cost recording from the internal systems of the supplier play a key role [9]. The factors for the value and risk assessments merge into the Customer Lifecycle Value. This sums up all incoming and outgoing payments within a customer lifecycle and reflects the potential of a customer for a supplier. The individual measurement of the customer lifecycle value represents the central strategic parameter for the decisions of pricing in the usage phase [24].

7. Optimize Pricing Model: The price metric provides the operational link between measurable KPIs and the price [9]. For this, an assessment basis as a logic of linking the individual KPIs with the price model plays a central role. An assessment basis for the data-based evaluation of the pricing model consists of systematised data types [13]. The pricing for availability and usage-based pricing models requires data from the machine, output-based price models require data from the performance process of the customer, and success-based price models require business data from the customer [23]. Following this, the design of mechanisms for pricing model optimisation must take place by identifying potential win-win situations by adjusting the price model. To determine improved price points for price elements, the expected additional value-in-use of the customer must be determined in monetary terms [21]. For this purpose, the financial effect of an improvement of an operational key figure of the customer must be determined and transferred to a single price unit of the price metric [3]. The pricing point is to be set that a supplier only gains a price premium from its customer when the KPIs increase [3]. For the implementation of mechanism for pricing model optimisation, the measurable added value for the customer and thus the theoretical willingness to pay cannot be fully skimmed in practice but only a part of the added value should be monetised [13].

8. Invoice Price Metrics: To record variable value-based price parameters, the active integration of customer-specific data is required to determine the level of the price. This requires the linkage of pricing processes from the supplier with the cyber-physical systems of the customer [9]. By using these systems, the supplier records data such as machine hours, output quantity or energy consumption in the use phase of the customer. To ensure a structured data acquisition for price calculation structures and processes must be created to transfer the data from the customer for the pricing processes of the supplier. The formula and the method for the calculation of the price is property of the price metric [6]. In the contract is the way how the customer pays for the solution and the frequency for the periodic payment defined. With periodic invoicing of the price for the offered service-system, the data-based charging of fixed and variable parameters is possible [13].

5. Conclusion

Manufacturing companies in the production industry increasingly provide their products, services, and smart services as integrated solutions within subscription business models. This enables suppliers to establish a close, long-term partnership within the usage phase of the machine. The objective is to facilitate a win-win situation for both the customer and the supplier, in which a customer receives an increase in value through continuous data-based optimisation of the offered solution, and the supplier receives continuous revenue from a customer that depends on the value of the customer. To achieve this, the charged price by the supplier must be related to the added value for the customer. The digital available data of the customer enable to form such a pricing based on value-in-use. However, most approaches for value-based pricing for the production industry focus on transaction-based business. This approach allows only an a priori assessment of the anticipated value. The actual value-in-use for the customer and therefore an incentive to the supplier to continuously improve the solution is not created. Accordingly, suppliers in practice do not use a structured process for pricing. As a result, suppliers lack knowledge regarding the way they should offer their solutions or fail to fully exploit the potential for optimising the value they provide when offering their solutions.

Therefore, this paper proposed a novel process model for pricing smart-product-service-systems based on the value-in-use of the customer for subscription business in production industry. Using the structured modelling approach, a model framework for subscription pricing is developed and subsequently equipped with concrete process steps and tasks within these process steps. The framework for this is an abstracted business process of pricing with the four elements reference basis, value assessment, price dimension, and price metric as well as two consecutive phases of the subscription business, the sales phase and the usage phase. To provide a practical method to an operator, a process model with eight consecutive process steps is derived within this framework based on the characteristics of subscription business models. Each of these process steps is assigned with concrete tasks and characteristics for the process specification. This allows suppliers to use the presented process model as a blueprint for developing new or adapting existing pricing processes for subscription business. The approach enables suppliers to establish a standardisable and successful long-term pricing for the subscription business. This offers opportunities for both suppliers and customers in the production industry to strengthen the competitiveness through close and profitable business relationships and to develop new growth potentials. From a scientific point of view, the approach offers a novel, operationally oriented process model for pricing, which enables value-in-use to be considered based on an a posteriori pricing.

References

[1] Schuh, G., Frank, J., Jussen, P., Rix, C., Harland, T., 2019. Monetizing Industry 4.0: Design Principles for Subscription Business in the Manufacturing Industry. 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC) (25), 1–9.

[2] Harland, T., 2018. Chancen und Risiken durch Subscription-Modelle im Maschinenbau. RWTH Aachen. Dissertationsvortrag, 10 December 2018, Aachen.

[3] Hermann, U., 2019. Digitalisierung im Industrieunternehmen: Die Chancen der digitalen Ökonomie der Dinge erkennen, entwickeln und erfolgreich umsetzen. Apprimus Verlag, Aachen.

[4] Mansard, M., Cagin, J.-M., 2019. Reaping the Recurring Benefits of Industry 4.0.: Mansard, Michael; Cagin, Jean-Michel (2019): Reaping the Recurring Benefits of Industry 4.0. A manufacturing executive playbook for business model transformation towards new revenue streams. Accessed 3 September 2019.

[5] Tzuo, T., Weisert, G., 2018. Subscribed: Why the subscription model will be your company's future - and what to do about it. Portfolio/Penguin, New York, New York, viii, 246.

[6] Lah, T., Wood, J.B., 2016. Technology-as-a-Service Playbook: How to Grow a Profitable Subscription Business. Point B Inc, s.l., 1363 pp.

[7] Truong, L., Polar, K., 2017. 4 Effective Subscription Pricing Strategies - Zuora. https://www.zuora.com/guides/4-effective-subscription-pricing-strategies/. Accessed 31 March 2021.

[8] Yamamoto, S., Sharma, P., 2019. Defining a Winning Subscription Pricing Model. https://www.zuora.com/guides/defining-a-winning-subscription-pricing-model/. Accessed 31 March 2021.

[9] Simon, H., Fassnacht, M., 2016. Preismanagement: Strategie - Analyse - Entscheidung - Umsetzung, 4., vollständig neu bearbeitete und erweiterte Auflage ed. Springer Gabler, Wiesbaden, 616 pp.

[10] Stoppel, E., 2016. Nutzungsabhängige Preissysteme auf industriellen Märkten. Springer Fachmedien Wiesbaden, Wiesbaden, 207 pp.

[11] Schönung, M.M., 2008. Kundenwertorientierte Preissetzung für Leistungssysteme im Maschinen- und Anlagenbau. Zugl.: Aachen, Techn. Hochsch., Diss., 2008. Apprimus, Aachen, 250 pp.

[12] Schuh, G., Gudergan, G., Kampker, A. (Eds.), 2016. Management industrieller Dienstleistungen: Handbuch Produktion und Management 8, 2. Aufl. 2016 ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 325 pp.

[13] Frohmann, F., 2018. Digitales Pricing: Strategische Preisbildung in der digitalen Wirtschaft mit dem 3-Level-Modell. Springer Gabler, Wiesbaden, 286 pp.

[14] Belz, C. (Ed.), 1997. Industrie als Dienstleister. Thexis, St. Gallen, 328 pp.

[15] Lusch, R.F., Vargo, S.L. (Eds.), 2014. The Service-dominant Logic of Marketing: Dialog, Debate, and Directions. Routledge, London, 449 pp.

[16] Macdonald, E.K., Kleinaltenkamp, M., Wilson, H.N., 2016. How Business Customers Judge Solutions: Solution Quality and Value in Use. Journal of Marketing 80 (3), 96–120.

[17] Haberfellner, R., 2019. Systems Engineering: Fundamentals and Applications. Birkhäuser, Cham.

[18] Dombrowski, U., Adams, T., Dreyer, M., 2017. Risikobasierte Preisgestaltung von Full-Service-Verträgen: Ein Vorghen zur Gestaltung von Full-Service-Verträgen am Beispiel eines Instandhaltungsunternehmens von Windkraftanlagen. ZWF 112 (11), 748–752.

[19] Hansen, M.S., 2018. How to build a subscription business: Guide, Inspiration and Cases, 2. Edition ed. Bookboon.com, Frederiksberg.

[20] Bliemel, F., Adolphs, K., 2003. Wertorientierte Preisstrategien, in: Diller, H., Herrmann, A. (Eds.), Handbuch Preispolitik. Strategien - Planung - Organisation - Umsetzung. Gabler Verlag, Wiesbaden, pp. 137–154.

[21] Keränen, J., Jalkala, A., 2013. Towards a framework of customer value assessment in B2B markets: An exloratory study. Industrial Marketing Management 42 (8), 1307–1317.

[22] Podratz, K., 2009. Prozess zur kundennutzenorientierten Preisfindung für industrielle Dienstleistungen. Unternehmen der Zukunft 10 (2), 60–62.

[23] Stoppel, E., Roth, S., 2017. The conceptualization of pricing schemes: From product-centric to customer-centric value approaches. Journal of Revenue and Pricing Management 16 (1), 76–90.

[24] Kunschert, M., 2009. Der Kundenwert im Industriegütermarketing. Kölner Wissenschaftsverlag, Köln, 316 pp.

[25] Tuli, K.R., Kohli, A.K., Bharadwaj, S.G., 2007. Rethinking Customer Solutions: From Product Bundles to Relational Processes. Journal of Marketing 71 (3), 1–17.

Biography

Tobias Leiting, M.Sc. M.Sc. (*1992) has been working as a project manager and PhD Candidate at the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2017. In his position, he leads industrial projects for the management and implementation of services and smart services. His research focuses on the marketing and pricing of smart services.

Prof. Dr.-Ing. Günther Schuh (*1958) has been director of the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2004. Prof. Dr.-Ing. Günther Schuh is director of the chair of production engineering at RWTH Aachen University. In addition, he was founder and CEO of the electrical car manufacturer e.GO Mobile.

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 the St. Gobain-Automotive Group for 10 years and lead the management of European plant logistics. In addition, he was responsible for the worldwide coordination of vehicle development projects.

Dr.-Ing. Jana Frank (*1990) has been head of service management at the Institute for Industrial Management (FIR) at the RWTH Aachen University since 2019. In her position she manages industrial projects and has the scientific lead for the service management department with the focus on subscription business management, service engineering, lean services and new work.

CONFERENCE ON PRODUCTION SYSTEMS AND LOGISTICS
CPSL 2021

2nd Conference on Production Systems and Logistics

Using Digitalization As An Enabler For Changeability In Production Systems In A Learning Factory Environment

Alexander Rokoss, Matthias Schmidt
Institute of Product and Process Innovation, Leuphana University of Lüneburg, 21335 Lüneburg, Germany

Abstract

Existing as well as new production systems need to adapt to meet increasing requirements of global, volatile markets. Companies aim to keep their production systems physically changeable while at the same time raise their level of digitalization. The human is a key aspect in the design process of production systems. Humans can only systematically develop suitable solutions if they have the necessary competencies to make reliable decisions. A promising approach to teach the required competencies is the use of gamification in form of learning factories. Approaches for learning factories addressing either digitalization or changeability are already described in literature. Although it lacks an examination on how to teach competencies in a learning factory environment, it enables participants to use digitalization methodically to promote changeability in production systems. This paper describes the required competencies regarding the implementation of changeability in production systems as well as competencies to support changeability making use of digitalization. In addition, it presents a concept for a learning factory workshop. Each round describes a new external impact on the production system that needs to be actively addressed by implementing interactive solutions containing digitalization as well as changeability methods.

Keywords

Learning factory; Gamification; Changeability; Digitalization; Competencies

1. Introduction

In the course of the digital transformation and increasingly volatile markets, companies have to digitalize their production systems while at the same time remain capable of change. A central point in the design of such systems are the professionals creating the production system, who can only systematically work out the creative process of the conception if they have the necessary competencies. One promising approach is the use of learning factories to teach the required competencies. These small-scale factories provide learners with the opportunity to actively take part in implementing and improving production processes in a realistic environment and to experience the results of their decisions without delay. The content learned in this way does not represent a mere extension of technical expertise but aims on developing competencies in dealing with methods and situations. These competencies can then also be applied on an industrial scale with little transfer effort [1]. Depending on the purpose and orientation, learning factories teach a very specific subarea of engineering, whereby the tasks to be solved are always derived from real problem situations and therefore equip learners with the tools to react immediately to such problems in real life situations [2].

The global digital transformation and the associated megatrends are creating a continuously changing market situation that is forcing companies to constantly adapt their production systems and, ideally, design them to be capable of change [2]. Varying customer requirements, changes in legislation and short-term strategical adjustments are examples of internal and external influences that force companies to transform their production systems constantly [3]. At the same time, digitalization is putting pressure on manufacturing companies. To operate efficiently, quickly and thus competitively, companies require digital support processes.

As mentioned above, professionals face the task of designing future production systems to be both changeable and digitally supported. To be able to do this, they need access to training that enables them to acquire the necessary skills and competencies. Literature shows that the focus of most learning factories, operated at corporate or university level, is on lean management workshops and demonstrating how digitalization technologies can be implemented to support lean manufacturing [5,4]. Although changeability has already been implemented in the form of modularized assembly workstations [6], literature on existing learning factories does not contain any examples that explicitly address the proactive design of changeability supported by digital transformation technologies. A description of a learning factory on changeability supported by digitalization technologies can help to improve the training of professionals who design future production systems.

The subsequent section describes the basic function of learning factories and the way of teaching competencies. Furthermore, it gives a brief overview on changeability in factories and digitalization indexing. It provides the framework for the analysis of the needed competencies to improve changeability in a factory, supported by digitalization technology in section three. Section four presents the learning factory workshop concept that has been developed. Lastly, the conclusions of the paper are summarized and future research possibilities are outlined.

2. Theoretical foundation on learning factories, changeability and digitalization

2.1 Learning Factories

Learning factories give participants the opportunity to actively engage in the implementation and improvement of production processes in a realistic environment while experiencing the results of their decisions without time delay [7]. Learning factories originate from the derivation of the principle of university clinics in the medical field [5] whereby both the high practical relevance of the students and the direct proximity of research to daily practice were taken into account. Although the term "learning factory" was first used in a research proposal by a consortium led by Penn State University [5] the literature agrees that the movement for the widespread use of learning factories originated in Central Europe, and from there expanded around the world [5,8]. A distinction is made between learning factories for research purposes and those for the education and training of professionals [5,4]. This paper considers only the teaching component. Depending on the purpose and orientation, learning factories teach a very specific part of engineering, whereby the tasks to be solved always derive from real world problems and therefore equip the learners with tools for immediate reaction [2]. As a result, learners develop professional competencies while remaining in a non-risky environment [9]. The approach of physically experiencing a learning process promises a significantly deeper anchoring of the imparted information as well as more joy in the learning process [11,10]. Another major advantage of learning factories is the fact that experimental learning is demonstrably more suitable for generating knowledge [12]. The lack of consequences of errors for ongoing business processes in the real environment clearly promotes curiosity and the willingness to experiment in learning factories. In their paper "Digital Learning Factories: Conceptualization, Review and Discussion" Haghighi et al. [10] break down four central areas of a learning factory in industrial engineering, on which this paper is based.

1. *Gaining an experience of different concepts and principles of production systems.*
2. *Gaining an experience regarding working with machines, tools, or physical equipment in the factory.*
3. *Obtaining social experiences such as group work and the ability to confront upcoming challenges.*
4. *Gaining an experience of working in realistic production environments and the meaning of situated cognition which is bodily involvement.*

The last two areas listed are sufficiently documented in literature in the general context of learning factories [6]. Therefore, this paper is limited to the principle of changeability in production systems and the use of digitalization technologies as tools needed to improve changeability.

The majority of existing learning factories currently address topics of lean management as well as how to deal with digital transformation technologies [5,7,13]. In a recently published study Martinez et al. show the rapid increase in publications dealing with learning factories since 2015 [6]. In the field of changeability these papers address for example the general handling of volatile market conditions [14], reconfigurable assembly processes [15] or building capabilities for agility in learning factory settings [16]. In the field of digitalization, the number of populations is larger. Lozano et al. examined 70 papers on digitalization and point out that 65 of them contain the term cyber-physical-system [17]. Common concepts and technologies such as RFID, digital milkrun, pick-by-light and how to teach the appropriate skills in using these technologies are already well documented [18]. In their paper "Digitalization as a catalyst for lean production: A learning factory approach for digital store floor management", Meissner et al. analyze the influence of digitalization on lean management topics, whereby adaptability is only mentioned briefly [19]. A more holistic approach is presented by Graz University of Technology in their LEAD (Lean, Energy efficient, Agile and Digital) learning factory. They describe the following requirements as prerequisites for building a learning factory for transformability: (1) Holistic case framework including a comprehensive image of the value chain, different types of external uncertainties and customer needs; (2) Access to ecosystem data; (3) IT systems (hardware/software) to process data; (4) Modular and flexible production setup and (5) Multiple case scenarios [14]. The combination of changeability and digitalization in a learning factory environment has not yet been deeply investigated. The authors used the value chain of the lean management learning factory already existent, allowing access to the ecosystem data and IT systems. Therefore, the concept in this paper focuses on changeable production setups and developing game round scenarios. As described below, some basics on changeability and digitalization are required to enable the creation of a tailored concept for a learning factory workshop.

2.2 Changeability

Changeability is defined as the ability to make short-term, but at the same time economic, technological, logistical, organizational and personnel changes that exceed the flexibility held in reserve, when the need arises [3]. The majority of changes themselves in a production system take place in the dimensions of the number of units, quality, time, product and cost structure [19]. Turbulences in the environment of a production system act as so-called change drivers, which are defined as internal or external influences that force a production system to change through their direct influence [3]. Examples of change drivers are: changing customer requirements, changes in legislation or internal strategy changes. The literature distinguishes between three classes of change drivers: target drivers, element drivers, and number-of-units and variant drivers, which have different effects on the production system [3]. These drivers are contrasted by the change enablers. The primary change enablers are universality, mobility, scalability, modularity and compatibility [20].

Universality describes the ability of man and machine to manufacture a wide variety of products and variants and their usage in changing tasks. Mobility describes the dynamic movability of objects within a factory.

The design of the objects allows them to be rearranged in a short time by means of rollers or by using an overhead crane. Scalability can be considered as the property of a production system to be able to expand and reduce its own capacities in spatial, technical and personnel terms. Modularity is understood as the ability of a production system to provide standardized units that can be exchanged for other, equally standardized parts as required. Finally, the change enabler compatibility creates the basis for networking individual units and devices within the production system by providing interfaces [20]. Meyer-Schwickerath et al. proposed a framework for the development and evaluation of response strategies to counter change drivers with the targeted use of change enablers [21]. It will be used later to identify certain competencies regarding changeability. Changeability per se, as well as the cited characteristics of change enablers, have so far been described only in terms of their physical characteristics. However, to adequately address constantly changing markets and the associated need for changeability, the digital aspects of change capability should also be considered [2].

2.3 Digitalization

The field of digitalization in learning factories is broad. For finding competencies to support the implementation of changeability in production systems, this paper takes into account two major publications. On the one hand, the concept is based on the industry 4.0 maturity index [22], on the other hand it considers a study that aims for identifying and categorizing the 25 most relevant competencies for digitalization [23]. In the latter paper competencies were categorized regarding information technology, multidisciplinary, electrical, business administration and mechanical engineering. An explanation of the identified competencies is provided in the next section.

Schuh et al. differentiate digitalization further and developed a model for classifying digitalization and industry 4.0 into so-called "benefit-oriented development stages" [22]. Stage 1 of the path describes the "computerization" of a company. Already widespread in most companies, computerization addresses the cost-effective, precise and low-error handling of routine processes, for example in the form of a computer-controlled milling machine. The second stage of the development path is "connectivity," which involves networking the machines that work in isolation as part of computerization. The Internet is also part of connectivity and enables centralized control of production facilities spread across the globe. "Visibility" forms the third stage of the model and enables the operator of a production facility to collect real-time data using numerous sensors and a collection network. The aggregation of the captured data allows the production facility to be mapped as a "digital shadow" in a virtual environment where the current operating status of each machine can be viewed. The fourth stage of the development path is the application of the digital shadow to better understand the complex interrelationships within the company and thus gain "transparency" about the processes taking place. Derived from the transparency of the data, forecasts can be calculated in the fifth stage of the model, which simulate different future scenarios by specifying a probability. The sixth and final stage of the development path addresses the "adaptability" of a production system. This refers to the autonomous reaction of the system to internal and external influences and the independent application of automated and self-optimizing counteractions [22].

3. Methodology to identify game rounds for the learning factory concept

Based on the existing structures, the learning factory concept in this paper is grounded on a well-timed production system that uses kanban for both material supply and order processing. However, to familiarize the participants with the full potential of digitalization technologies the production system works completely analog. The desired form of the learning factory is a production system, which, in accordance with the benefit-oriented development path, records all relevant production data, gains transparency of the processes from the data obtained and finally generates or independently initiates recommendations for action. In this way, the system can respond early to drivers of change and is designed in a way that drivers can be countered within a very short time by changing the production system itself.

To structure the conceptual design of the learning factory, the teaching objectives were defined based on the first two areas of interest referring to Haghighi et al. as mentioned earlier. In the area of changeability, the didactic objectives are the secure handling of the concepts of flexibility and changeability as well as the control loop of changeability according to Nyhuis [24]. They should have an understanding of different change drivers and their consequences for the system within the turbulent environment and the interactions between change drivers and change enablers.

In the field of digitalization, the aim of the learning factory is to provide participants with a comprehensive overview of the technologies underlying the concept. Building on this theoretical foundation, the aim of the learning factory is to emphasize sufficiently the importance of connecting the available technologies with reference to each individual case. In addition, participants should have an understanding of the respective areas of application of the technologies and the actions required to use them effectively and efficiently. Lastly, a key point is to sensitize the participants to the possibility of a technical failure in digitally supported systems. Ideally, at the end of the learning factory workshop, participants will have a brief knowledge of the interaction between existing changeability and the use of digitalization technologies, both negatively and positively. Ultimately, the learning factory needs to enable the participants to independently design adaptable, digitally supported production systems or to improve existing systems through targeted actions.

In the course of game rounds carried out in the learning factory, the participants will encounter different drivers of change that either make it necessary to use the implemented changeability or force the players to spontaneously transform the existing system. Optimally, they will then resort to the available digitalization technologies and use them as supporting change enablers to master the drivers.

Therefore, in the conceptual design of the learning factory it is advisable to interpret the process backwards. Identifying the change enablers and their digital supporters serves as a starting point to derive the necessary competencies to handle them. This is followed by filtering from a catalog of change drivers that require the broadest possible spectrum of changeability competencies in order to be mastered. The first step is to record competencies relating to the general implementation and integration of changeability in production systems. The aim of the analysis is to create a competency catalog based on the matrix structure for evaluating technical-methodical competencies according to Abel et al. [9]. The second step is to identify competencies regarding the five primary change enablers as well as the three change driver classes enhancing the competency catalog. Step three is to include the competencies for exploiting the potential of digitalization in the context of the benefit-oriented development stages by Schuh et al. Next, a fictitious scenario is assigned to each potential change driver in a real-world environment, which could have an effect on the production system within the learning factory environment. Subsequently, the change driver scenarios are examined, evaluated and classified as applicable or non-applicable with regard to their feasibility in a highly simplified model environment as used in the learning factory. This is followed by the identification of four different entry scenarios of change drivers on the production system, which are examined as game round scenarios along the digitalization development path, which is needed to be able to recognize the change at an early stage in the game round. In preparation for the dynamics of participants, possible options for action for each changeability scenario are being identified.

In accordance with the procedure described above, the following section develops a catalog of competencies for changeability supported by digitalization in a learning factory environment.

4. Competencies on changeability and digitalization

For the documentation and description of competencies relevant to the professional field, the matrix structure developed by Abel et al. for the assessment of professional-methodical competencies is being used [9]. It subdivides a main competence (e.g. ability to perform a method) into several sub-competencies. Each sub-competence is assigned the respective action, e.g. analysis of the actual process. Furthermore, required professional knowledge and underlying conceptual knowledge is assigned to each sub-competence. As an example: for the application of the Single Minute Exchange of Die (SMED) method, the knowledge of the concept and the knowledge of the purpose and area of application of the SMED method would be required accordingly as professional knowledge. In addition, process knowledge is required on how the resulting shorter setup times affect material stocks and throughput times. Finally, as conceptual knowledge, there needs to be the understanding that setup optimization primarily makes sense if the associated equipment represents the bottleneck in the production system [9].

For the determination of the necessary sub-competencies in the field of changeability, the four design fields technology, logistics, organization and human resources were used [19]. Furthermore, the identified sub-competencies are based on the procedure for developing and evaluating response strategies [21]. Table 1 shows an example tuple from the competence catalog. The full catalog consists of 51 sub-competencies, categorized underneath Table 1.

Table 1: Listing example from the competence catalog

No.	Partial competence	Action	Professional knowledge	Process knowledge	Conceptual knowledge
1	Participants have the ability to identify the need for change	Evaluate production system in terms of existing adaptability.	Knowledge of the concept and general relationship between the production system and the environment.	Knowledge of the implications of sufficient or insufficient transformability.	Understanding that transformability only needs to be improved if it is not sufficient for potential drivers.

A total of 35 sub-competencies in the field of changeability were identified and described in accordance with the matrix structure for assessing technical and methodological competencies. The sub-competencies were subdivided into the three competencies "Assessing changeability" (10 items), "Recognizing change drivers" (5 items) and "Using change enablers" (20 items). The sub-competencies for dealing with change enablers were further subdivided into the five enablers universality (4 items), mobility (4 items), scalability (3 items), modularity (5 items) and compatibility (4 items).

The digitalization competency catalog consists of competencies relating to the classification and application of technologies in the context of the digitalization maturity index [22] as well as the 25 competencies most relevant for digitalization [23]. Latter were reduced to the five most relevant sub-competencies in the field of digitalization by means of a pairwise comparison. These include the fields of identification systems, data analytics, cloud computing, databases and network technology, from each of which a sub-competence was derived (5 items). Based on the Industry 4.0 maturity index, the six levels of the model Computerization (1 item), Connectivity (2 items), Visibility (2 items), Transparency (2 items), Predictability (2 items) and Adaptability (2 items) result in a total of 11 sub-competencies.

The catalog of competencies to be taught contains a total of 51 sub-competencies, broken down into the main competencies of handling changeability, change drivers and change enablers, supported by digitalization.

5. Learning factory on digitally supported changeability in production systems

For the implementation in a learning factory concept, it is necessary to bring the taught competencies into a logical sequence so that game events are understandable to the participants. Therefore, in the present case, it is necessary to integrate the concept of changeability into the game rounds, considering possible driver categories. These drivers then need to be countered with a specific change enabler and, in parallel, it must be communicated how the different levels of the digitalization maturity index can help to recognize and initiate the necessary change.

For the pre-selection of possible change driver scenarios, the exemplary change driver catalog according to Klemke et al. [25] is used for the preliminary selection of possible change driver scenarios. It distinguishes 16 changeability drivers in eight categories being legislator and associations, customer and market, suppliers, employees, technology, competitors and miscellaneous. Each category gets assigned two exemplary change drivers, which are subdivided into ten element drivers, five number-of-units and variant drivers and one target driver. To address the reference to possible instabilities within the production systems, the failure of a digitalization technology is also included in the catalog as an element driver of the "Technology" category, resulting in a total of 17 change drivers.

To examine the listed change drivers for their possible use in the context of the learning factory, each scenario is assigned an estimated time frame in which the driver will presumably influence the production system. Additionally, a scenario will be given that would be suitable for integration in the learning factory. The gradation of the assigned change timeframes is subdivided into less than two months, two to twelve months, twelve to 60 months, and timeframes of more than 60 months.

Grouping the individual stages of the digitalization maturity index enables the entire development path to be demonstrated, starting from a completely analog production system in the first game round. In each of the remaining three rounds of the game, two stages of the development path are combined and implemented. The first game round only serves to convey the mechanism of changeability and socialization with the existing analog production system. Table 2 shows the structure of the identified game rounds and the respective driver scenarios within the rounds.

Table 2: Structure of the identified game rounds

Round	**Benefit-oriented development path**	**Change driver**	**Change driver scenario**	**Change Enabler**	**Digitalization actions**	**Improvement actions**
1	Not digitized	Changed part quality	Only philips screwdriver is available, participants face slotted screws because no others are available.	Modularity	-	Universal screwdriver
2	Computerization & Connectivity	New products from competitors Change in make-or-buy strategy	Competition has individual license plates, initially sales slump, but then goes up again when individual license plates can also be attached.	Scalability Universality	Participants write small program to enable 3D printer to print rims on a conveyer for mass production. The implement their program into the MES software.	Use production machine rims (3D printer) for license plates Reduce employees first, then increase again
3	Visibility & Transparency	New/changed requirements for product properties	SUV is requested by customers, requires different undercarriage chassis. Customer requests more transparency on traceability.	Universality Mobility	Participants develop a KPI system by writing SQL-Queries that are directly displayed on their andon board while producing.	Layout rearrangement Design vehicle base universally Move customer order decoupling point
4	Predictability & Adaptability	Sales figures slump Shorten delivery time	New competitor on the market Company has to shrink healthily or switch to pure chassis production in a very short time, in order to survive.	Compatibility Scalability	Participants create a simple timeseries-based machine learning software that predicts customer demand on certain car variants.	Layout rearrangement Use forecast on customer orders are more interested in other vehicles, but the chassis for the SUV is very popular, but must be manufactured with shorter delivery times

Depending on the round in the workshop, participants have to apply different change enablers to continue to operate the production system economically. Additionally, they have to implement a digitalization approach, in order to fulfill costumer demands. The driver of the wrong screws in round 1 can be countered by means of modularity, in which different bits can be added to the modular screwdriver. As this round only serves the purpose of teaching the basic principle of changeability, there will be no digitalization implementations. Round two involves converting the existing production system and using the existing 3D printer to produce license plates. To do so, they have to write a basic javascript software snippet, that connects the printer to the MES system in order to automate rim printing. As the sale related key performance indicators (KPIs) suddenly drop, participants must first reduce and then increase the staff employed by means of scaling. In round three, participants need to change the entire product design to create a modular base that is fitted with the respective chassis for normal vehicle and off-road vehicle at the end of the assembly process. The resulting change in layout is initially not possible due to bolted down workstations. It must be possible to solve this problem by mobilizing the tables, for example by means of rollers [8]. In order to meet customer requirements regarding the transparency of the production and shipping progress, the group has to extract some KPIs from the MES. They do this by completing a short course on structured language queries (SQL) that enables them to build their own KPI-System from the MES data. KPIs are now displayed on an andon board throughout the whole game round, enabling the group to anticipate the shift in customer variant demands. Round four starts with the participants following up on how they determined the point of change in round three, based on their KPI system. Instructors provide the group with examples of supervised learning approaches based on free software to predict certain outcomes. Participants customize the given models according to their experiences from the game and the developed KPI system. The goal is to develop a system that predicts changes in customer demands based on data from the latter round. The driver of collapsing sales related KPIs, detected early by the new prediction system, leaves the participants with two options. They can either scale down the production system to produce the remaining vehicles in a more economical way or convert the entire production system to a chassis-only production.

The aim of the game rounds is for the participants to perceive not only the basic mechanisms of changeability but also how the use of digitalization buys them time to be able to recognize and initiate a necessary change in reasonable time. To point out the weak points of these technologies, one of the technologies should be eliminated in one of the last two rounds to make the participants aware of their dependence on the overall function of the implemented IT systems.

6. Conclusions and Outlook

Designing changeable, digitally supported production systems requires certain competencies of the planning staff. Learning factories can help to acquire these competencies without influencing live production operations. For this purpose, the respective learning factories must be precisely tailored to the topics to be taught to enable the participants to gain the necessary competencies while playing. Existing learning factories address either changeability or digitalization. Systematic analysis of the field of changeability in combination with the stage model of the digitalization and industry 4.0 maturity index helps to understand the interactions. In order to be able to convey the leverage effect of digitalization technologies on the early detection and initiation of a change in a production system, a multi-round learning factory workshop is required. The process described in this paper represents an approach to communicating the topics of changeability, change drivers and change enablers in the context of digitalization. Follow-up tasks are to continue implementing the concept in practice, measuring the learning successes, and iteratively improve the learning factory concept. Additionally, the concept could be enhanced by integrating new technology like machine learning.

References

[1] Abele, E., Metternich, J., Tisch, M., 2019. Learning Factories. Springer International Publishing, Cham.

[2] 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). Herbert Utz Verlage, München, 72 pp.

[3] Wiendahl, H.-P., ElMaraghy, 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.

[4] Blöchl, S.J., Schneider, M., 2016. Simulation Game for Intelligent Production Logistics – The PuLL® Learning Factory. Procedia CIRP 54, 130–135.

[5] Abele, E., Metternich, J., Tisch, M., Chryssolouris, G., Sihn, W., Elmaraghy, H., Hummel, V., Ranz, F., 2015. Learning Factories for Research, Education, and Training. Procedia CIRP 32, 1–6.

[6] Martinez, P., Vargas-Martinez, A., Roman-Flores, A., Ahmad, R., 2020. A science mapping study on learning factories research. Procedia Manufacturing 45, 84–89.

[7] Sackey, S.M., Bester, A., Adams, D., 2017. INDUSTRY 4.0 LEARNING FACTORY DIDACTIC DESIGN PARAMETERS FOR INDUSTRIAL ENGINEERING EDUCATION IN SOUTH AFRICA. SAJIE 28 (1).

[8] Matt, D.T., Rauch, E., Dallasega, P., 2014. Mini-factory – A Learning Factory Concept for Students and Small and Medium Sized Enterprises. Procedia CIRP 17, 178–183.

[9] Abel, M., Czajkowski, S., Faatz, L., Metternich, J., Tenberg, R., 2013. Kompetenzorientiertes Curriculum für Lernfabriken. Werkstattstechnik online : wt, Springer VDI Verlag, Düsseldorf 103 (3), 240–245.

[10] Haghighi, A., Zadeh, N.S., Sivard, G., Lundholm, T., Eriksson, Y., 2014. Digital Learning Factories: Conceptualization, Review and Discussion, Göteborg.

[11] Deslauriers, L., Schelew, E., Wieman, C., 2011. Improved Learning in a Large-Enrollment Physics Class. Science 332 (6031), 862–864.

[12] Cachay, J., Wennemer, J., Abele, E., Tenberg, R., 2012. Study on Action-Oriented Learning with a Learning Factory Approach. Procedia - Social and Behavioral Sciences 55, 1144–1153.

[13] Veza, I., Gjeldum, N., Mladineo, M., Celar, S., Peko, I., Cotic, M., Ljumovic, P., Stojkic, Z., 2017. Development of Assembly Systems in Lean Learning Factory at the University of Split. Procedia Manufacturing 9, 49–56.

[14] Karre, H., Hammer, M., Ramsauer, C., 2018. Learn how to cope with volatility in operations at Graz University of Technology's LEAD Factory. Procedia Manufacturing 23, 15–20.

[15] Ahmad, R., Masse, C., Jituri, S., Doucette, J., Mertiny, P., 2018. Alberta Learning Factory for training reconfigurable assembly process value stream mapping. Procedia Manufacturing 23, 237–242.

[16] Karre, H., Hammer, M., Ramsauer, C., 2019. Building capabilities for agility in a learning factory setting. Procedia Manufacturing 31, 60–65.

[17] Lozano, C.V., Vijayan, K.K., 2020. Literature review on Cyber Physical Systems Design. Procedia Manufacturing 45, 295–300.

[18] Hulla, M., Hammer, M., Karre, H., Ramsauer, C., 2019. A case study based digitalization training for learning factories. Procedia Manufacturing 31, 169–174.

[19] Nyhuis, P. (Ed.), 2010. Wandlungsfähige Produktionssysteme. GITO-Verl., Berlin, 468 pp.

[20] Nyhuis, P., Reinhart, G., Abele, E. (Eds.), 2008. Wandlungsfähige Produktionssysteme: Heute die Industrie von morgen gestalten. Technische Informationsbibliothek u. Universitätsbibliothek; PZH Produktionstechnisches Zentrum, Hannover, Garbsen, 72 pp.

[21] Meyer-Schwickerath, B., Schuch, L., Rahman, O.A., Delitzsch, S., 2013. Wandlungsfähigkeit von KMU erhöhen: Eine anwendungsorientierte Vorgehensweise zur Entwicklung und Bewertung von Reaktionsstrategien. Werkstattstechnik online : wt, Springer VDI Verlag, Düsseldorf 103 (7), 624–630.

[22] Schuh, G., Anderl, R., Gausemeier, J., Hompel, M. ten, Wahlslter, W., 2017. Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten. Deutsche Akademie der Technikwissenschaften.

[23] BiBB, 2015. Next generation competencies for a digital world. BWP – Berufsbil-dung in Wissenschaft und Praxis (6), 33–35.

[24] Nyhuis, P., 2010. Wandlungsfähige Produktionssysteme. Gito.

[25] Klemke, T., Mersmann, T., Wagner, C., Goßmann, D., Nyhuis, P., 2011. Bewertung und Gestaltung der Wandlungsfähigkeit von Produktionssystemen. ZWF 106 (12), 222–227.

Biography

Alexander Rokoss (*1993) studied industrial engineering at the Leuphana University of Lüneburg. In the context of his master studies, he developed the software applications for the Leuphana Learning Factory. Since 2018, he works as a research associate in the field of machine learning in production management at the Institute of Product and Process Innovation (PPI) at Leuphana.

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.

Conference On Production Systems And Logistics
CPSL 2021

2nd Conference on Production Systems and Logistics

Multi-Sensor Identification Of Unmarked Piece Goods

Marc Ungen[1,2], Louis Louw[2], Daniel Palm[1]
[1]*Reutlingen University, Reutlingen, Germany*
[2]*Stellenbosch University, Stellenbosch, South Africa*

Abstract

The seamless fusion of the virtual world of information with the real physical world of things is considered the key for mastering the increasing complexity of production networks in the context of Industry 4.0. This fusion, widely referred to as the Internet of Things (IoT), is primarily enabled through the use of automatic identification (Auto-ID) technologies as an interface between the two worlds. Existing Auto-ID technologies almost exclusively rely on artificial features or identifiers that are attached to an object for the sole purpose of identification. In fact, using artificial features for the purpose of identification causes additional efforts and is not even always applicable. This paper, therefore, follows an approach of using multiple natural object features defined by the technical product information from computer-aided design (CAD) models for direct identification. By extending optical instance-level 3D-Object recognition by means of additional non-optical sensors, a multi-sensor automatic identification system (AIS) is realised, capable of identifying unpackaged piece goods without the need for artificial identifiers. While the implementation of a prototype confirms the feasibility of the approach, first experiments show improved accuracy and distinctiveness in identification compared to optical instance-level 3D-Object recognition. This paper aims to introduce the concept of multi-sensor identification and to present the prototype multi-sensor AIS.

Keywords

3D-Object recognition; Automatic identification; Computer-aided design (CAD); Direct identification; Multi-sensor identification; Natural identifiers

1. Introduction

Digitalisation and Industry 4.0 are far-reaching fields of action with great relevance for society, education and economy in all areas of life worldwide. The core of the ideas behind Industry 4.0 results from research and development activities in engineering or natural sciences which are closely linked to information and communication technology as well as automation technology [1]. One promising concept behind Industry 4.0 is the seamless integration of the real world of things with the virtual world of information, widely referred to as the 'Internet of Things' (IoT) [2]. Tracing back the origins of this term, it can be found that it was already coined around the year 2000 by the founders of the original Massachusetts Institute of Technology (MIT) Auto-ID Center, which is nowadays part of the Auto-ID Labs research network [3]. The latter institutions conducted research on automatic identification (Auto-ID) technologies for industry aiming to establish the foundation for the IoT [2], and they identified Auto-ID technologies as an important technology for the future advancement of the IoT and therefore Industry 4.0.

2. Motivation

In the context of industry, the term 'automatic identification technology' refers to a variety of techniques for the purpose of logistical goods identification within material flow systems, aiming to synchronise the information flow with the material flow by collecting identity information [4, 5]. Besides less known techniques, visual code identification (approx. 70%) and radio-frequency identification (RFID) dominate today's applications [5]. Common to both techniques is the use of artificial features for the purpose of identification. Visual codes are applied to objects either by labelling (e.g., thermal transfer labels, etc.) or by direct marking (e.g., laser engraving, printing, etc.) [5-7]. For identification using RFID, active or passive electronic transponders are attached to objects, which are available in various designs (e.g., adhesive labels, press-in cartridges, etc.). Active RFID transponders differ from passive ones by an integrated power supply in form of a battery [7].

It is obvious that the utilisation of the above-mentioned artificial identifiers requires additional process steps for their attachment to objects. If an object itself is to bear such an artificial identification feature, its geometry must also be appropriate in size and shape, which must already be considered in its design phase and offers difficulties in application if the geometry cannot be adapted as is the case with functional surfaces for example. When going through different production steps, it may be necessary to remove previously applied artificial identifiers and afterwards reapply them to objects as they would be destroyed (e.g., painting, thermal treatment, etc.), also offering drawbacks in application. Active RFID transponders furthermore require maintenance in case their batteries have to be replaced. In conclusion, the use of artificial identifiers generates costly additional efforts and is not always applicable.

Identification based on features that characterise objects by nature is referred to as 'direct identification' in the literature [5-8]. Natural identification features of objects are mass properties, geometry, surface structure/texture, colour/appearance and material [5-10]. Using these features instead of artificial identifiers eliminates the issues discussed in the previous paragraph.

Modern machine vision (MV) systems can recognise three-dimensional objects from image data based on their appearance or geometry, known as '3D-Object recognition' [11]. As a basis for the recognition, computer-aided design (CAD) models can be used, which serve to define known objects within a knowledge base [12]. However, these MV systems are only capable of interpreting visually perceptible characteristics, which limits their suitability for direct identification as they cannot distinguish based on the remaining natural identification features. As a consequence, the distinctiveness and accuracy of identification using 3D-Object recognition are limited, which also limits possible application tasks.

In industry, CAD models are widely used as virtual product models containing related product information [13]. In particular, the technical information, as a subset of the product information, provide details on the natural characteristics that can be used for identification purposes mentioned above [14]. CAD models are thus an ideal source for direct identification, defining industrial objects throughout the entire life cycle following the paradigm of product data management (PDM).

This leads to the core idea behind multi-sensor identification, which consists of extending the perceptual capacity of MV systems for 3D-Object recognition by using further sensors to detect the natural identifiers not considered so far. The resulting multi-sensor automatic identification system (AIS) is thus no longer dependent on the use of artificial identification features, while at the same time offering higher potential use for identification tasks than MV systems.

3. State of the art

This section briefly summarises the state of the art in related fields and provides insight into the theoretical foundations of multi-sensor identification.

3.1 Existing approaches for direct identification

While there are many approaches or technologies for indirect identification by means of artificial identification features, there are only a few approaches for direct identification by exploiting the natural features of objects, often referred to as 'fingerprints' [9].

Laser surface authentication (LSA) takes advantage of naturally and randomly occurring imperfections on the surfaces of objects to be identified. These imperfections cause diffuse scattering when exposed to a focused laser beam. This diffuse scattering, also known as laser granulation or laser speckle, captured by photodetectors arranged at different angles generates a pattern of the reflected intensity unique to the object. Applying statistical methods, a binary descriptor can be generated from this intensity pattern, which can then be used for identification similar to a human fingerprint [9, 15].

The characterization of grinding imprints by means of their roughness profile offers another possibility for direct identification. Due to the grain structure of grinding wheels and the wear occurring during grinding processes, random grinding patterns are created on the surfaces of workpieces. After recording such a ground surface with a high-resolution camera, descriptors can be generated by means of various algorithms. These descriptors can be used for identification and are also robust to perturbations like corrosion [10].

3.2 Instance level 3D-Object recognition based on CAD models.

The recognition of three-dimensional objects from image data is commonly referred to as '3D-Object recognition'. In the literature, a basic distinction is made in terms of the level of recognition and the type of input image data used. Instance level recognition describes the identification of distinct object instances [11]. In contrast to category-level recognition, this means that an object to be recognized can be explicitly assigned to a known object within the recognition knowledge base. For the generation of this recognition knowledge base, which describes the objects known to the recognition system, CAD models can be used [12]. Numerous feature descriptors are available for both 2D and 3D image data. Due to the fact that the depth information in 3D image data provides higher quality, 3D feature descriptors are preferably used for recognition based on CAD models [11, 12, 16]. Most modern 3D feature descriptors encode surface normal information, that is obtained from point clouds [16, 17]. The development of such descriptors has been stimulated by the increasing availability and affordability of 3D sensors, that directly capture point clouds of objects.

CAD models primarily describe the geometric shape of objects and appear in three different basic forms of representation: wireframe models, surface models and solid models [18]. By means of virtual rendering or sampling methods, these representations can be transformed into point clouds.

For describing complete point clouds of objects or CAD models, global descriptors are particularly suitable. One very accurate and performant global descriptor is the clustered viewpoint feature histogram (CVFH) descriptor [16]. Global descriptors are the basis of global processing pipelines for 3D-Object recognition. The basic global recognition pipeline consists of the steps 'description' and 'matching', optionally six degrees of freedom (6DoF) pose estimation can be performed. In a preceding offline process, descriptors are generated for all point clouds of CAD models and saved as a recognition knowledgebase. The actual recognition takes place in an online process, which commences with describing the sensor point cloud of an object to be identified in order to generate a so-called scene descriptor. The latter is then matched against all descriptors of known objects within the recognition knowledgebase, where the quality of each match is assessed by means of a distance metric. If there is a sufficient degree of agreement between the scene and the knowledge base descriptors, an object is recognised. By performing the 6DoF pose estimation step, the three-dimensional geometric transformation that transfers the CAD point cloud to the sensor point cloud can be determined [19].

3.3 Technical information within product models

Product models in the industry comprise defining information for products. One subcategory of this product information is technical information, typically originating from product development. In particular, the geometrical and technological information within the technical information contain details regarding the physical features of products or objects [14].

CAD models are product models which are used in particular to store and distribute technical information [18]. There are many different data formats for CAD models, which differ greatly regarding their information content. Typically, CAD models contain the following technical information: geometries, mass properties, material properties. Some CAD models can serve as a basis for photorealistic rendering, which generates detailed information on their appearance or texture.

4. Concept for multi-sensor AIS

This section presents the generic process for identification based on natural identification features, a software processing chain based on it as well as a hardware concept for a multi-sensor AIS.

4.1 Generic process of identification

In the literature, there are only a few formulations regarding the process of identification (e.g. [20]). These formulations mainly focus on the superficial phases that are passed through to identify an object in a practical environment while neglecting the process of identification itself. For this reason, the formulation of a generic process of identification is necessary.

Identification essentially consists of the assignment of identification features of an object to be identified to the identification features of a known object to retrieve associated identity information. The fundamental prerequisite for performing an identification is, therefore, an identification knowledge base (e.g., register of citizens) that defines individual identities by linking identification features (e.g., appearance from passport photo, height, eye colour, etc.) with identity information (e.g., name, address, etc.). Identity is therefore only valid within the framework of an identification knowledge base and can have different levels of uniqueness. The level of uniqueness of identity depends on the uniqueness of the identification features used. If several objects possess identical identification features (e.g., article number), they share the same identity (e.g., master data record) within the scope of the identification knowledge base. The identification features must therefore always be chosen with regard to the desired identification task and the required level of uniqueness.

An identification knowledge base K consists of j pairs $\left(S_{F_j}, S_{I_j}\right)$ which define identities (see Equation 1). S_F and S_I describe sets of identification features and identity information with arbitrary length (see Equation 2 and Equation 3).

$$K = \left\{\left(S_{F_1}, S_{I_1}\right)_1, \left(S_{F_2}, S_{I_2}\right)_2, \dots, \left(S_{F_j}, S_{I_j}\right)_j\right\} \quad (1)$$

$$S_F = \{F_1, F_2, \dots, F_k\} \quad (2)$$

$$S_I = \{I_1, I_2, \dots, I_l\} \quad (3)$$

The actual process of identification now describes the search for the pair (S_F, S_I) within K, where S_F matches the feature set S_f of an object to be identified (see Equation 4).

$$S_f = \{f_1, f_2, \dots, f_m\} \quad (4)$$

The above formulation of the generic process of identification is the basis for the multi-sensor identification procedure as several features are evaluated. This is not the case with state-of-the-art identification systems that rely on only one identification feature (e.g., label with identification number). In addition to the uniqueness of each individual identifier, a higher degree of uniqueness is created through the way several identification features are combined.

4.2 Processing chain for multi-sensor AIS based on CAD-Models.

The processing chain consists of an offline and an online process. The offline process generates the identification knowledge base from the CAD models, while the online process performs the actual identification.

4.2.1 Offline process for identification knowledge base generation

The inputs for the offline process (see Figure 1) are the CAD models of objects to be identified. For each CAD model, a descriptor for 3D-Object recognition is created after transformation into a point cloud. Together with the identity information (identification number and description), as well as the information on weight, colour/appearance and centre of mass location, these descriptors are stored in the identification knowledge base. The identification knowledge base is the result of the offline process and is subsequently provided to the online process.

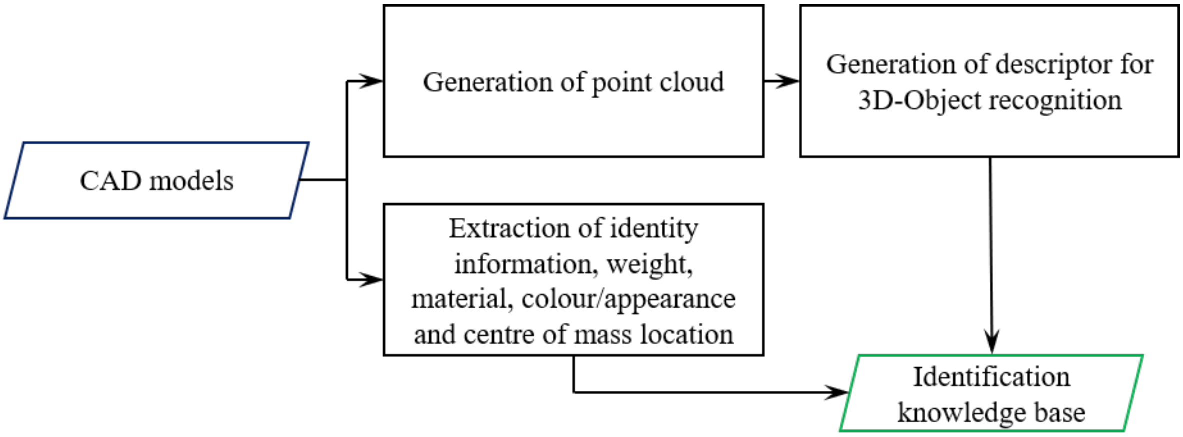


Figure 1: Offline process for identification knowledge base generation from CAD models; input (blue) output (green)

4.2.2 Online process for multi-sensor identification

The inputs for the online process (see Figure 2) are sensor inputs acquired from a scene containing an object to be identified as well as the identification knowledge base from the offline process. The inputs "Sensor point cloud" and "Sensor colour/appearance" can be obtained from 3D-Scanners. Regarding the inputs "Sensor weight", "Sensor material" and "Sensor centre of mass location" a sensor platform can be used which is presented at a later stage.

From the sensor point cloud, a descriptor for 3D-Object recognition is generated. Based on weight, material, and colour/appearance information the identification knowledgebase is prefiltered in order to only pass matching candidate objects to the "3D-Object recognition and 6DoF pose estimation" step. The descriptors of these candidate objects point clouds from the prefiltered identification knowledge base are then matched against the descriptor of the sensor point cloud in the sense of 3D-Object recognition. For the recognised object, an estimation of the 6DoF pose is carried out, which yields the geometric transformation in order to transform the CAD point cloud into the sensor point cloud. By means of this transformation, a "Centre of mass validation" step can be performed. This step serves to compare the centre of mass location from the CAD model with the centre of mass location detected by sensors, aiming to further improve the accuracy of identification. After successfully performing all steps of the online process, the identity information originating from the identification knowledge base are available.

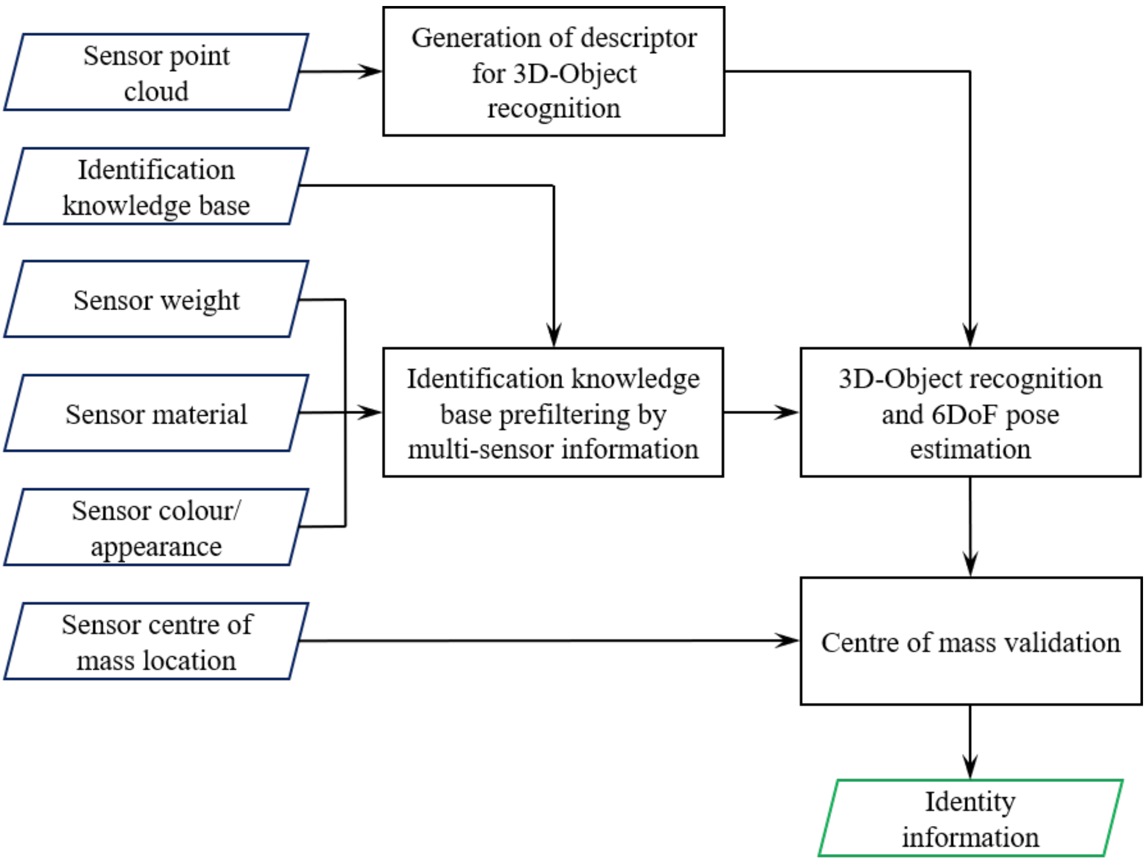


Figure 2: Online process for multi-sensor identification; inputs (blue) outputs (green)

4.3 Hardware concept for multi-sensor identification

The concept for the acquisition of sensor inputs for the offline process consists of a structured-light 3D-Scanner in combination with a sensor platform (see Figure 3). The 3D-Scanner captures the point cloud and colour/appearance of the identification object, which is settled on the rotary device of the sensor platform. The rotary device rotates the object to be identified during scanning in order to obtain a complete scan. An inductive sensor is used to detect the object's material, which makes it possible to distinguish between metals and non-metals. Furthermore, the sensor platform consists of a weighing plate, which uses four force sensors to determine the weight and centre of mass location of the object.

As Figure 3 indicates, three coordinate systems occur in this arrangement: The coordinate system of the 3D-Scanner (indexed "S"), the coordinate system of the sensor platform (indexed "SP") and the coordinate system of the object (indexed "O"). These coordinate systems are important for the "Centre of mass validation" step of the online identification process as geometric transformations have to be applied in order to compare the centre of mass location between the CAD model and the physical identification object.

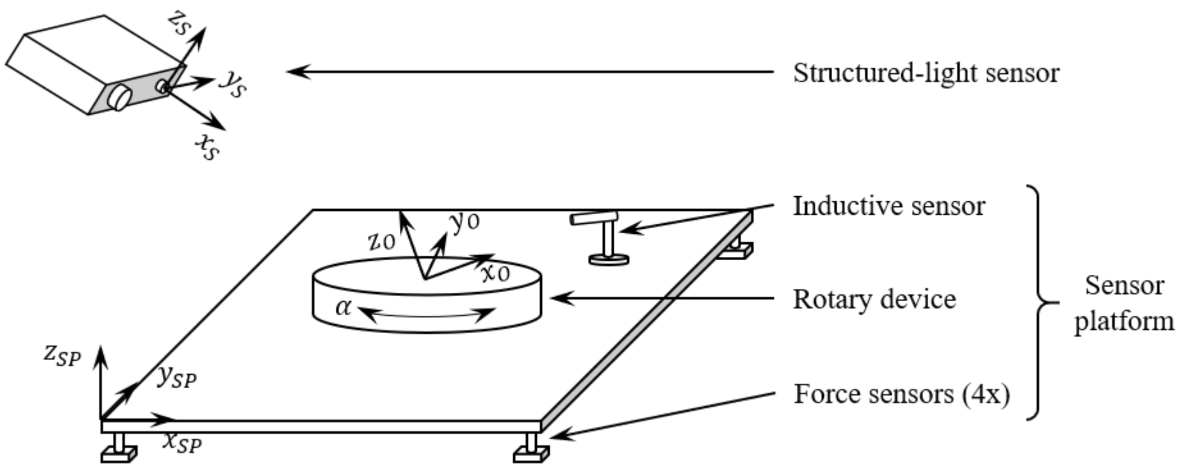


Figure 3: Sensor hardware concept for multi-sensor identification

5. Prototype implementation of multi-sensor AIS

Based on the above concept, the prototype for multi-sensor identification was implemented (see Figure 4). An ‘HP 3D Structured Light Scanner Pro S3’ in combination with an ‘HP Turntable Pro’ were used for the structure light sensor and the rotary device. The sensor platform is a custom implementation based on four load cells with HX711 amplifiers and an inductive proximity switch connected to an Arduino Nano. The weighing plate with the force sensors attached to it is made of steel so that a 3D-printed mount for the inductive proximity switch can be slid onto it by means of magnets.

Since 3D-Object recognition is computationally intensive, the above hardware was connected to a computer, serving as the main processing device. The software for the multi-sensor identification processing chain (see Section 4.2) was implemented using the programming languages Python and C++. In particular, the module performing 3D-Object recognition was implemented using the Point Cloud Library (PCL) written in C++ and then integrated into Python via the Pybind11 library.

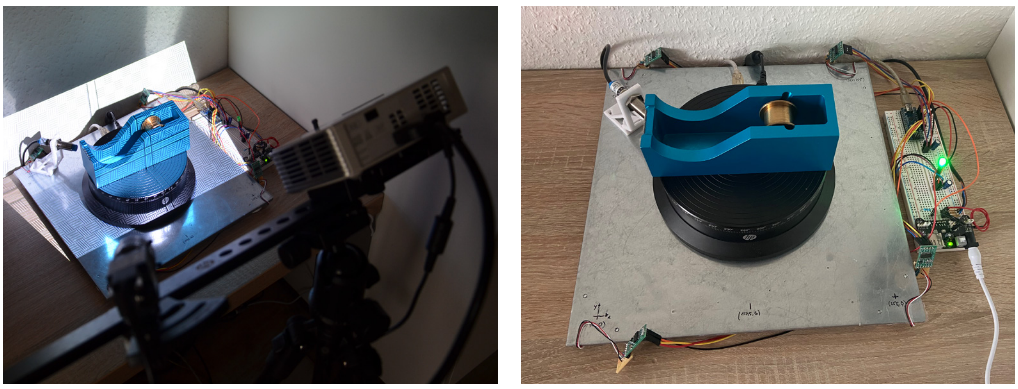


Figure 4: Prototype multi-sensor AIS with an identification object settled on the rotary device

6. Experimental verification of prototype multi-sensor AIS

For the experimental verification of the multi-sensor AIS, a set of 21 objects was compiled. Attention was paid to the fact that the objects vary in geometry, colour/appearance, weight, centre of mass location and material. From the set 16 objects were soda cans, differing in geometry (Ø67mm*115mm, Ø53mm*135mm, Ø67mm*115mm), colour/appearance (manufacturer specific design) and weight (full/empty). These soda cans were selected to be difficult to distinguish, as sometimes only one identifying feature differs from another object.

Using the prototype, 63 multi-sensor identifications were carried out, all of which were successful. The identification features were recorded reliably in all cases, leading to the appropriate identity information each time. On average, it took 3:08 minutes to identify one object by means of the online process. The main limiting factor here is the time needed for scanning, which is approximately 80% of the time span.

As a result of the investigation, it can be stated that the implemented prototype and algorithm for multi-sensor identification works well with the selected objects. The distinct identification of optically (geometry, colour/appearance) indistinguishable objects was always successful, which represents a clear advantage over conventional instance-level 3D-Object recognition systems in terms of accuracy and distinctiveness. There is still a lot of potential for more efficient software implementation and better hardware setup in order to drastically reduce the time needed for the identification process.

7. Summary and outlook

This paper introduces the concept of multi-sensor identification. Based on a newly introduced generic process of identification (see Section 4.1), a processing chain for multi-sensor identification using CAD models (see Section 4.2) is presented. A hardware concept for collecting multi-sensor information as inputs for the processing chain is described and implemented as the multi-sensor AIS prototype. Results of experimental verification proof the feasibility and effectiveness of multi-sensor identification.

Further research is needed in order to improve the hardware concept and the processing chain. The use of a robotic arm with integrated sensors in the gripper for collecting multi-sensor information and the manipulation of objects to be identified in front of a 3D-Scanner is worth investigating. Furthermore, the presented approach could be combined with the existing methods for direct identification (see Section 3.1) to utilize these distinct identification features. Here, CAD models would be useful for defining the location of 'fingerprints' (see Section 3.1) on objects. Since some types of 3D scanners already have integrated high-resolution cameras, using these identification features would not need any additional hardware.

References

[1] Deckert, R., 2019. Digitalisierung und Industrie 4.0: Technologischer Wandel und individuelle Weiterentwicklung. Springer Gabler, Wiesbaden.

[2] Uckelmann, D., Harrison, M., Michahelles, F., 2011. An Architectural Approach Towards the Future Internet of Things, in: Uckelmann, D. (Ed.), Architecting the Internet of Things, 1 ed. Springer, Heidelberg, pp. 1-24.

[3] Santucci, G., 2010. The internet of things: Between the revolution of the internet and the metamorphosis of objects, in: Sundmaeker, H. (Ed.), Vision and Challenges for Realising the Internet of Things, 1 ed. European Commission – Information Society and Media DG, Brussels, pp. 11-24.

[4] Jünemann, R., Beyer, A., 1998. Steuerung von Materialfluß- und Logistiksytemen: Informations- und Steuerungssyteme, Automatisierungstechnik, 2 ed. Springer, Berlin.

[5] Ten Hompel, M., Büchter, H., Franzke, U., 2008. Identifikationssysteme und Automatisierung. Springer, Berlin.

[6] Hippenmeyer, H., Moosmann, T., 2016. Automatische Identifikation für Industrie 4.0. Springer Vieweg, Berlin.

[7] Finkenzeller, K., 2015. RFID-Handbuch: Grundlagen und praktische Anwendungen von Transpondern, kontaktlosen Chipkarten und NFC, 7 ed. Carl Hanser, München.

[8] Krämer, K., 2002. Automatisierung in Materialfluss und Logistik: Ebenen, Informationslogistik, Identifikationssysteme, intelligente Geräte, 2 ed. Deutscher Universitäts Verlag, Wiesbaden.

[9] Buchanan, J., Cowburn, R., Jausovec, A., Petit, D., Seem, P., Xiong, G., Atkinson, D., Fenton, K., Allwood, D., Bryan, M., 2005. 'Fingerprinting' documents and packaging, in: Nature, 436, p. 475.

[10] Dragon, R., Mörke, T., Rosenhahn, B., Ostermann, J., 2011. Fingerprints for machines - Characterization and Optical Identification of Grinding Imprints, in: Mester, R. (Ed.), Pattern Recognition. Springer, Berlin, pp. 276-285.

[11] Alhamzi, K., Elmogy, M., Barakat, S., 2014. 3D Object Recognition Based on Image Features: A Survey, in: International Journal of Computer and Information Technology, 03(03), pp. 651-660.

[12] Aldoma, A., Vincze, M., Blodow, N., Gossow, D., Gedikli, S., Rusu, R., Bradski, G., 2011. CAD-Model Recognition and 6DOF Pose Estimation Using 3D Cues, in: IEEE International Conference on Computer Vision Workshops. IEEE, Piscataway, pp. 585-592.

[13] Hehenberger, P., 2020. Computergestützte Produktion: Eine kompakte Einführung, 2 ed. Springer, Berlin.

[14] Eigner, M., 2014. Technische Organisation des Produktentwicklungsprozesses, in: Eigner, M., Roubanov, D., Zafirov, R. (Ed.), Modellbasierte virtuelle Produktentwicklung, 1 ed. Springer Vieweg, Berlin, pp. 227-266.

[15] Cowburn, R., 2008. Laser surface authentication – reading Nature's own security code, in: Contemporary Physics, 49(5), pp. 331-342.

[16] Mateo, C., Gil, P., Torres, F., 2014. A Performance Evaluation of Surface Normals-based Descriptors for Recognition of Using CAD-Models, in: 11th International Conference on Informatics in Control, Automation and Robotics. IEEE, Piscataway, pp. 428-435.

[17] Carvalho, L., von Wagenheim, A., Torres, F., 2019. 3D object recognition and classification: a systematic literature review, in: Pattern Analysis and Applications, 22. Springer, London, pp. 1243-1292.

[18] Roubanov, D., 2014. Mechanikkonstruktion (M-CAD), in: Eigner, M., Roubanov, D., Zafirov, R. (Ed.), Modellbasierte virtuelle Produktentwicklung, 1 ed. Springer Vieweg, Berlin, pp. 115-136.

[19] Aldoma, A., Marton, Z., Tombari, F., Wohlkinger, W., Potthast, C., Zeisl, B., Rusu, R., Gedikli, S., Vincze, M., 2012. Point Cloud Library: Three-Dimensional Object Recognition and 6DoF Pose Estimation, in: IEEE Robotics & Automation Magazine, 19(3). IEEE, Piscataway, pp. 80-91.

[20] Lolling, A., 2003. Identifikationssyteme in der Logistik: Übersicht und praxisorientierte Auswahl. Huss-Verlag, München.

Biography

Marc Ungen (*1993) studied the double-degree master programme Digital Industrial Management and Engineering at Reutlingen University and Stellenbosch University (2019-2021). Before that, he studied mechanical engineering at Reutlingen University (2014-2018).

Louis Louw (*1974) has been an associate professor at Stellenbosch University since 2018. Previously Prof. Dr.-Ing. 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) studied mechanical engineering at the University of Stuttgart and received his doctorate from the Technical University of Vienna. He was appointed Professor of Supply Chain Management at the ESB Business School at Reutlingen University in 2013. He is head of the Research Centre for Value Creation and Logistics Systems and the Reutlingen Center for Industry 4.0.

2nd Conference on Production Systems and Logistics

Pre-Selection Of Suitable Regression Methods For The Determination Of Interactions And Forecasts In Global Production Networks

Günther Schuh[1], Andreas Gützlaff[1],
Julian Ays[1], Niklas Rodemann[1], Miha Podbreznik[1]
[1]*Laboratory for Machine Tools and Production Engineering WZL of RWTH Aachen University, 52074 Aachen, Germany*

Abstract

The locations of many manufacturing companies are distributed globally. This has led to the development of historically grown global production networks whose structure is often very complex, not transparent and influenced by many factors. The high number, as well as the volatility of the influencing factors and dependencies in the network additionally, complicate the network configuration. As a result, adaptation needs and optimization possibilities are recognized too late or not at all. In order to enable early recognition of saving potentials, active monitoring and analysis of changes and dependencies of the influencing factors on the production network is needed. The necessary consideration of a multitude of influencing factors requires further tools to be manageable by the network planner. Therefore, databased methods can be used as support for the forecast and the determination of dependencies of influencing factors. In other research fields, regression analysis is an established method for a databased analysis. This paper focuses on the use of regression analysis in global production networks. It is essential for an accurate analysis, to choose the right regression method out of the many different types in existence. A systematic literature review is conducted to establish an overview of regression methods used in other research fields. A search strategy is developed and implemented and the key findings of the literature review are derived and evaluated. In the second step, a new approach for the pre-selection of suitable regression methods for the determination of interactions and forecasts in global production networks is proposed.

Keywords

Regression Analysis; Global Production Network; Influencing Factors; Network Design

1. Introduction

Globally located production sites are interconnected with material, information and financial flows [1,2]. The challenge for companies is to manage their interconnected network of locations effectively and efficiently [2]. The growing complexity of global production networks complicates the management of the network and its quick adaptation to changes [3]. To achieve transparency and subsequently to overcome the complexity, an understanding of the network is of importance. Global production networks are characterized and defined by a multitude of influencing factors [3]. These influencing factors include internal data (e.g. production cost, production capacity, etc.) and external data (e.g. changes in local taxes, inflation, competitor's action, etc.) [4,5]. With a better understanding and forecast of influencing factors, predictions about the whole network can be made. This is why data-based analysis in global production networks is gaining importance [6]. A possibility for a data-based forecast of influencing factors is the use of regression analysis. Regression analysis is a data mining class [7], used for not only predictions and forecasts, but also the determination of causal relations between dependent and independent variables [8]. The distinctive differentiation between regression methods is in linearity [7,9,10]. Regression is widely researched and implemented in other research areas, e.g., it has been used in forecast models for economic and financial models [11], as evidence in court cases and as support for legislation [8]. It is the most widely used statistical approach in medical research [12]. COE ET AL. present the parallels between global production networks and the global financial network and propose an integration of finance into the global production network research [13]. Because of its characteristics and a proven track record in other research areas, this paper investigates which regression methods can be used and how to select a suitable regression method in global production networks. Regression methods can then be used to determine the interactions and forecasts in the network, e.g., forecasts of costs or determination of interdependencies between production sites.

2. State of the art

This chapter presents the state of the art approaches in the determination of interactions and forecasts in global production networks. Until now, the main influencing factors have already been identified and described. ABELE ET AL. identify and describe the main influencing factors in global production networks [14]. Many other authors develop catalogues of influencing factors such as NEUNER [15]. In this context, some authors focus specifically on methodologies for the network configuration. UDE develops a decision support method for the configuration and evaluation of global production networks [16]. REUTER ET AL. develop a method to evaluate performance differences between manufacturing sites, the correlations between key performance indicators and site characteristics are analysed [17]. Some publications propose and introduce databased analysis and support for decision making in global production networks. VERHAELEN ET AL. develop a methodology to support product allocation in global production networks, with the use of different data mining methods [18]. SCHUH ET AL. present a software tool for the configuration and optimization of global production networks [19]. GÖLZER ET AL. investigate the configuration and design of global production networks using Big Data [20]. MOURTZIS ET AL. propose the configuration of global production networks based on smart decision making [21]. HOCHDÖRFER ET AL. use a clustering analysis on product portfolios for the reduction of the planning complexity in global production networks [22]. MOSER ET AL. approach flexible migration planning in global production networks with the use of the Markov decision process [23]. A further step with the use of regression analysis in global production networks was made by RITTSTIEG [24]. The author investigates the cause-effect relationships of a few influencing factors with multivariant and univariant linear regression [24]. The results are not compared to other regression methods and the author does not offer a selection methodology for different regression methods [24]. TREBER presents the increase of transparency in production networks with the improvement of disruption management through increased exchange of information [25]. For the analysis, the author uses three different regression models [25]. In conclusion, the literature review indicates that there is no scientific publication investigating the potentials and the use of regression analysis in global production networks. Neither is there an established method for the selection of suitable regression models in global production networks. From numerous applications in other fields, it will be investigated which regression methods have potentials for use in global production networks.

3. Systematic analysis of suitable regression methods

To answer the research question, a systematic literature review methodology according to BRAMER ET AL. is used [26]. Systematic literature review is a method, that is effective, when used to answer a broad research question, because it represents a proven tool for summarizing many primary studies on a predefined object of investigation [26,27].

As described in the methodology, according to BRAMER ET AL., firstly a search strategy was developed, then tested, optimized, and finally reiterated [26]. Every step of the systematic search strategy was documented. As the database for the search, Web of Science was selected, since its database is, in comparison to other databases, more extensive in the fields of engineering and natural science [28]. In the next step, base search keywords were defined, expanded with synonyms, and connected with Boolean operators. To achieve the most accurate thematic fit and to limit the number of possible results, only the abstracts of the publications were used in the search. The search term was tested and optimized with multiple iterations. The final search term was:

$$AB = (regression\ AND\ (analysis\ \ OR\ method *)\ AND\ (forecast * \\ OR\ prognosis\ \ OR\ interrelation))$$

In the search, no timespan was defined, and all publication years were considered. The number of search results before filtering was 33.521.

A filtration and refinement process was defined (see Figure 1). In the first search, the used parameters were open access of publications and English as the language of the publications. After this step, the number of search results was reduced to 14.558. In the second step the results were filtered by the following Web of Science categories to achieve a thematic fit: “mathematics interdisciplinary applications”, “computer science artificial intelligence”, “computer science interdisciplinary applications”, “operations research management science”, “computer science information systems”. The initial results were evaluated, and the number of results was 343 publications, with the average citation of the publication cited being 7,43. To further reduce the number of the results and to ensure a high impact-factor, the publications were filtered according to the number of citations. The number of citations in the final selection is 7 overall citations or higher. This yielded 97 results. The list of publications was checked for errors and one publication was eliminated from the evaluation, because of unavailability. After the filtration and elimination process, the final number of evaluated publications is 96 (see Figure 1).

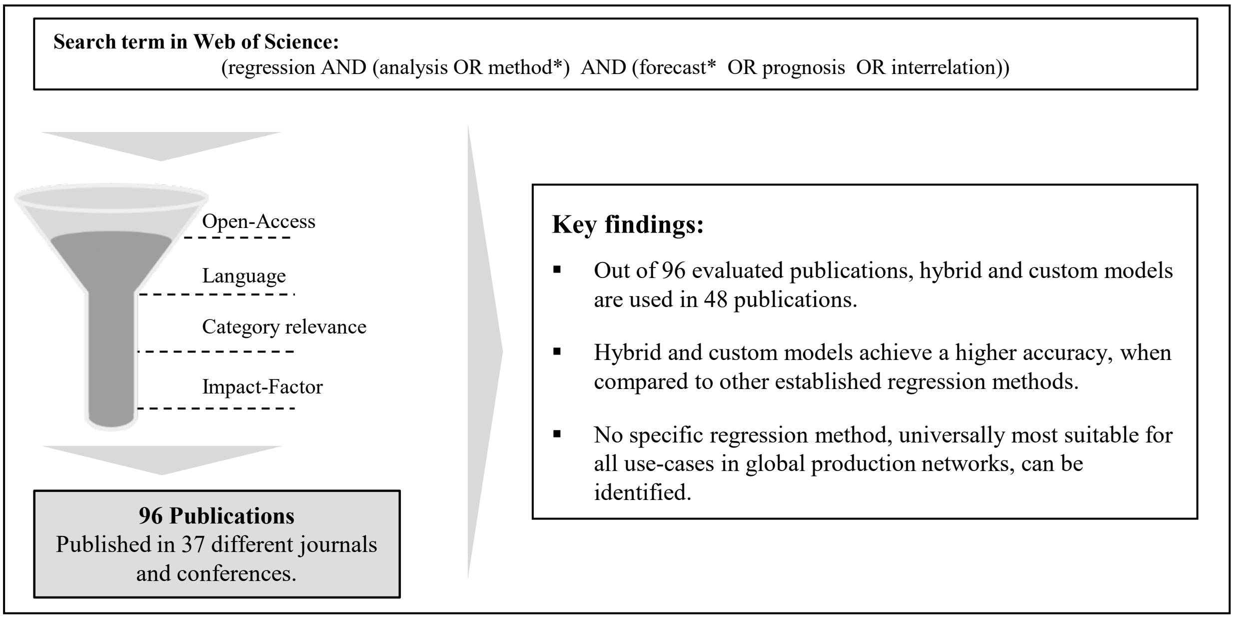


Figure 1: Overview of the systematic literature review

The publication year, of evaluated publications, ranges from the year 2003 to 2020. The total sum of citations is 2202 and the average citation is 22.94 per publication, ranging from 7 citations to 206 citations. The publications were published in 37 different journals and conferences.

Literature evaluation

The 96 evaluated publications mostly focus on the implementation of the models and methods. One publication discusses how to choose the best forecast model. Other 95 publication either describe the development of a new model or present the implementation of a model. Figure 2 shows how many times a model is used in the evaluated publications, in some, multiple methods were used and presented. In the following section, methods and models used in more than one reviewed publication are presented and a few use-cases are highlighted.

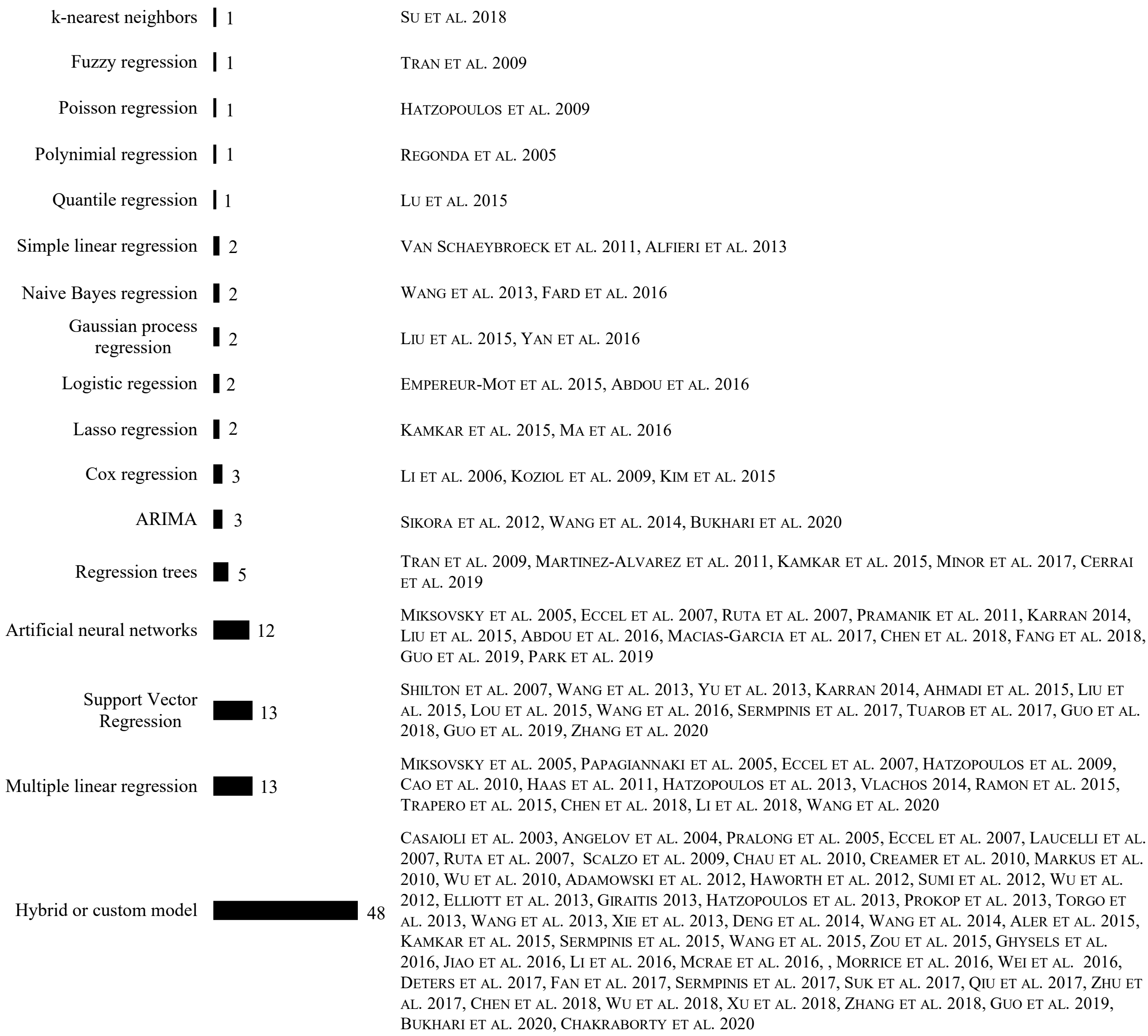


Figure 2: Summarized number of times the regression methods used in evaluated publications (N=96)

Hybrid or Custom Model

Hybrid or custom models are used in 44 publications. A hybrid model using both artificial neural networks (ANN) and support vector regression (SVR) proved a significant improvement in comparison to the use of a conventional ANN in a case for daily rainfall forecasting [29]. The hybrid model achieved the best result with filtering with single spectrum analysis. It was shown, that enumeration was more effective than correlation analysis since it can consider the nonlinear dependence between inputs and outputs [29].

In financial market forecasting, a new hybrid model was developed by combining autoregressive fractional integrated moving average (ARFIMA) and long-short term memory (LSTM) [11]. LSTM networks are a special kind of artificial neural networks. The prediction accuracy of the hybrid model was tested against autoregressive integrated moving average (ARIMA), ARFIMA and generalized regression neural network (GRNN) independently and showed 80% accuracy improvement [11]. With the high accuracy and reliability shown in the reviewed publications, hybrid and custom models offer the possibility to analyse any use-case in global production networks, since the model is adapted to the specific problem and is highly accurate.

Multiple and simple linear regression

A simple regression model is a linear regression model with a single response variable y and a single independent variable x. Multiple linear regression is an expansion of the simple linear regression and is used when more than one independent variable is given. [30]

A general multivariate regression model is a linear model consisting of k independent variables and n response variables, it can be represented using a matrix [31]. Linear regression can offer a simple approach to the analysis of influencing factors in global production networks, when dealing with simple and homogenous linear datasets. Linear methods are generally more susceptible to outliers but can be made more robust by regularization parameters [10].

Support Vector Regression

The support vector machines can be generalized to be used for regression. SVR is effective in real-value function estimation. It is a supervised learning method and it uses a symmetrical loss function, which equally adjusts high and low misestimates. The margin of tolerance is symmetrically placed around the function [32]. SVR have an excellent generalization capability with high accuracy of forecasts [32]. Due to its characteristics, it could be used to analyse influencing factors in global production networks with a complex and inhomogeneous dataset.

Artificial Neural Networks

The development of artificial neural networks was motivated by the information processing of the human brain [30]. An artificial neural network consists of neurons and connectors, which build a network. The classes of networks are mainly distinguished by the different network topologies and connection types, such as single-layer, multilayer, feed-forward or feed-back networks [30]. RUTH ET AL. use a robust ensemble of neural network regressors with smoothing of the output signal [33]. This composite of neural networks was tested against others in multiple competitions. Artificial neural networks are a possible method for the analysis of complex influencing factors and large datasets in global production networks.

Regression Trees

The regression tree can be considered a type of tree-based algorithm, with a continuous response variable [30]. In a publication, the M5P algorithm is used to create regression trees for improved earthquake prediction [34]. The M5P algorithm is a binary regression tree model where the last nodes are the linear regression functions that can produce continuous numerical attributes. It is a further development of the M5 algorithm introduced by QUINLAN [35]. This tree algorithm can handle high dimensionality tasks. The regression trees method could offer a use case in the event prediction of influencing factors in global production networks.

ARIMA

The autoregressive integrated moving average (ARIMA) model is mostly used for time series forecasting models. WANG ET. AL use the ARIMA model for precipitation simulations [36]. ARIMA models are widely used to calculate monthly time series with yearly variations. The authors slightly modify the model to account for inter-monthly variations, which are, as the authors suggest, often ignored. This modification improves the accuracy by 21%. It is concluded, that the ARIMA approach can be further improved with the use of neural networks and support vector machines [36]. This approach could be used for the forecast of influencing factors in global production networks since the model is fitted to time series data to better understand the data or to forecast data points.

Cox Regression

The Cox proportional-hazards model is a regression model, commonly used in medical research for the investigation of the association between the time, a specified event takes to happen, and other variables. The Cox regression approach has drawbacks since it does not enable the identification of interactions between variables [37]. KIM ET AL. propose an improved framework as an improvement to the traditional Cox model [38]. Although most of this method's use cases are in the field of medical research, it could, with some modifications, be used in global production networks.

Lasso Regression

Least absolute shrinkage and selection operator (Lasso) is a regression method commonly used to model the predicted risk of a likely outcome. It has often been used in areas, where the number of potential predictors is large relative to the number of observations [39]. Lasso regression has been shown to outperform many standard regression methods in some settings. MA ET AL. introduce a new framework to forecast retail sales using a multistage Lasso regression [40]. The improved framework results in significantly higher accuracy. This method could be of use in global production networks when large influencing factor constructs are used.

Logistic Regression

The logistic regression uses a logistic function to model a binary output variable [41]. The logistic regression model uses logistic regression to predict the odds of the binary outcome [41]. EMPEREUR-MOT ET AL. use logistic regression to calculate activity probabilities in clinical virtual screening methods [42]. In global production networks, this method could be used, for the odds prediction of event occurrence.

Gaussian Process Regression

Gaussian process regression is used, to profit from the normal distribution. Its property as a machine learning method allows automatic model building based on observations. The result is a probability distribution of possible interpolation functions and the solution with the highest probability [43]. A Gaussian process regression model has been introduced for forecasting the time series of wind energy [44]. The possibility of the forecast of influencing factors makes it a potential method for use in global production networks.

Naive Bayes Regression

Naive Bayes is usually used for classification tasks and is often more reliable than more sophisticated classification methods [45]. Naïve Bayes can also be used for regression. FARD ET AL. introduce new models with high accuracy for event prediction in longitudinal data [46]. Due to the high accuracy shown in the reviewed publications, naive Bayes regression could, when adapted, be used for the analysis of influencing factors in global production networks.

Key Findings

The key findings presented in Figure 1 were derived from the review of the publications. Other research areas use different regression methods for different use-cases and specific data sets. There is no specific method, that prevails in the number of uses. Hybrid and custom models are used in 48 publications. These differentiate themself from each other in the combination of models used and the structure of the model as well as the result expected. The highest accuracy achieved models, which were adapted and customized to a specific use-case or problem, because of their adaptation to the specific dataset and the expected result. Hybrid and custom models, which were compared in reviewed publications with other established methods, achieved a higher forecast and prediction accuracy. Because of different data sets, different influencing factors and different use-cases in global production networks, a universally best method for use in global production networks cannot be identified. Due to the different data sets of influencing factors, different regression methods could be most suitable for the analysis, with hybrid and custom models being the most accurate and most effort-intensive to implement.

4. Selection of suitable regression methods

As identified in section 3, there is no universally best regression method in global production networks. The analysed methods are suitable for different use cases depending on the data basis and the result requested. As accuracy is not the only criterion to be considered, a structured approach to the selection of a suitable model should be implemented.

This paper proposes a novel methodical approach to the pre-selection of regression methods (see Figure 3). The pre-selected regression methods, reviewed in the systematic literature review, are used as support.

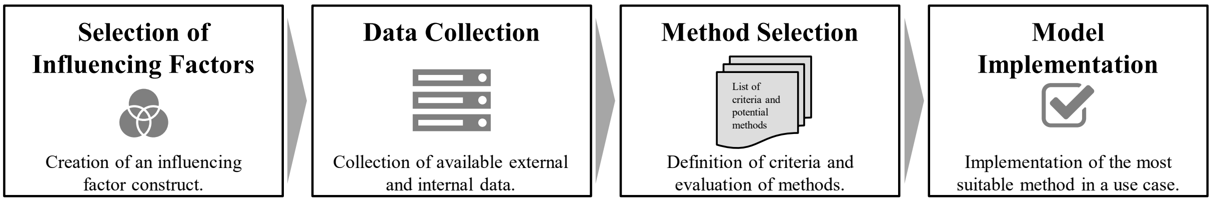


Figure 3: Proposed methodology for the method pre-selection

In the first step, all available internal and external data should be collected needed to describe a construct consisting of influencing factors, which were selected beforehand. Subsequently, the data is screened and reviewed and afterwards prepared for the analysis. If the collected data is homogenous, and there are little changes and outliers in data, and errors in forecasts are of little importance, a simple linear method or a method most convenient for the user can be chosen. Otherwise, a list of potential methods should be created and evaluated and structurally judged by experts according to criteria for a specific use-case. The criteria of YOKUMA ET AL. can be used as a guideline. YOKUMA ET AL. analyse two studies, where expert opinion, on the criteria for the selection of forecasting models, was examined and summarized in the following ranked list (from most to least significant): accuracy, cost savings from improved decisions, ease of interpretation, flexibility, ease of using available data, ease of use, quickly provided forecast [47].

Although accuracy is listed as the most important in the expert opinion, other criteria are nevertheless relevant and should be considered together in the selection and the development of new methods and models. For example, the aspects of convenience, market popularity, corporate guidelines and relative track record can be considered [48]. The criteria for the selection should be specific, weighted and can differentiate from use-case to use-case. In the creation of the list of potential methods, section 3 of this paper can be used as a guide, since these are the most relevant methods in other research areas in recent years. The best-rated model can be used, if the decisions made according to the analysis do not critical and standard model is sufficient. Otherwise, a new custom algorithm should be developed and tested against already established methods for accuracy and other selected criteria. The hybrid or custom models, used in reviewed publications, achieved, when compared to other algorithms, the highest accuracy. Potentially the highest accuracy in global production networks can be achieved with such a model. In the last step, the selected method should be implemented. The result of the implementation are forecasts and the relations between dependent and independent variables in the network.

5. Discussion and further research

The findings from the systematic literature review suggest, that the highest accuracy for regression models is achieved with hybrid or custom models, which are developed for a specific problem. The publications, that used standard regression methods and models mostly modified them and improved them for the specific use-case. One of the key findings is that there is not a universally accurate and best regression method in global production networks. The accuracy depends on many variables. One of them is available data, the more data there is available, the more accurate the forecast can be [49]. The newly developed and modified hybrid and custom models were validated with the standard regression methods and achieved a significant improvement of accuracy. As a next step, a new approach to the selection of regression methods in global production networks is proposed.

To apply regression methods in the context of global production networks, further research on the validation of the proposed methodology should be conducted. The influencing factors that have the most impact on global production networks have to be identified, characterized and defined. As a next step, a data mining model based on different regression analysis models for global production networks should be developed and implemented.

Acknowledgements

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

References

[1] Schuh, G., Prote, J.-P., Gützlaff, A., Rodemann, N., 2019. Gestaltung von Produktionsnetzwerken: Verknüpfung von datengetriebenen Analysen mit Expertenwissen zur Konfiguration globaler Produktionsnetzwerke. ZWF 114 (1-2), 7–10.

[2] Arndt, T., Lemmerer, C., Biegler, C., Sihn, W., Lanza, G., 2017. Steuerung globaler Produktionsnetzwerke: Entwicklung eines Standortrollenmodells zur dynamischen Bewertung von Gestaltungsmaßnahmen. wt Wekstattstechnik online 107 (4), 241–246.

[3] Váncza, J., 2016. Production Networks, in: Produ, T.I.A.f., Laperrière, L., Reinhart, G. (Eds.), CIRP Encyclopedia of Production Engineering. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–8.

[4] Ferdows, K., 2018. Keeping up with growing complexity of managing global operations. IJOPM 38 (2), 390–402.

[5] 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), 823–841.

[6] Kwon, O., Lee, N., Shin, B., 2014. Data quality management, data usage experience and acquisition intention of big data analytics. International Journal of Information Management 34 (3), 387–394.

[7] Cios, K.J., Pedrycz, W., Swiniarski, R.W., Kurgan, L.A., 2007. Data mining: A knowledge discovery approach. Springer, New York, NY, 606 pp.

[8] Sykes, A.O., 1993. An Introduction to Regression Analysis. Coase-Sandor Institute for Law & Economics (20).

[9] Otte, R., Otte, V., Kaiser, V., 2004. Data Mining für die industrielle Praxis. Hanser, München, 389 pp.

[10] Runkler, T.A., 2010. Data Mining: Methoden und Algorithmen intelligenter Datenanalyse. Vieweg + Teubner, Wiesbaden.

[11] Bukhari, A.H., Raja, M.A.Z., Sulaiman, M., Islam, S., Shoaib, M., Kumam, P., 2020. Fractional Neuro-Sequential ARFIMA-LSTM for Financial Market Forecasting. IEEE Access 8, 71326–71338.

[12] Jain, R.K., 1985. Ridge regression and its application to medical data. Computers and Biomedical Research 18 (4), 363–368.

[13] Coe, N.M., Lai, K.P.Y., Wójcik, D., 2014. Integrating Finance into Global Production Networks. Regional Studies 48 (5), 761–777.

[14] Abele, E., Dervisopoulos, M., Liebeck, T., 2006. Herausforderungen globaler Produktionsnetzwerke: Wie Unternehmen ihr Produktionsnetzwerk optimal aufstellen. wt Werkstattstechnik online 96 (4), 219–225.

[15] Neuner, C., 2009. Konfiguration internationaler Produktionsnetzwerke unter Berücksichtigung von Unsicherheit. Zugl.: Bayreuth, Univ., Diss., 2009, 1. Aufl. ed. Gabler, Wiesbaden, 316 pp.

[16] Ude, J., 2010. Entscheidungsunterstützung für die Konfiguration globaler Wertschöpfungsnetzwerke: Ein Bewertungsansatz unter Berücksichtigung multikriterieller Zielsysteme, Dynamik und Unsicherheit, 201 pp.

[17] Reuter, C., Prote, J.-P., Stöwer, M., 2016. Evaluation of Performance Differences between Manufacturing Sites. Procedia CIRP 50, 342–347.

[18] Verhaelen, B., Thomas, K., Häfner, B., Lanza, G., Schuh, G., 2017. Potenziale datenbasierter Produktallokationen, in: Wunder, T. (Ed.), CSR und Strategisches Management, vol. 114. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 96–100.

[19] Schuh, G., Potente, T., Kupke, D., Varandani, R., Hausberg, C., 2012. An Evolutionary Approach for Global Production Network Optimisation. Procedia CIRP 3, 382–387.

[20] Gölzer, P., Simon, L., Cato, P., Amberg, M., 2015. Designing Global Manufacturing Networks Using Big Data. Procedia CIRP 33, 191–196.

[21] Mourtzis, D., Doukas, M., Psarommatis, F., 2015. A toolbox for the design, planning and operation of manufacturing networks in a mass customisation environment. Journal of Manufacturing Systems 36, 274–286.

[22] Hochdörfer, J., Laule, C., Lanza, G., 2017. Product Variety Management Using Data-Mining Methods – Reducing Planning Complexity by Applying Clustering Analysis on Product Portfolios, 593–597.

[23] Moser, E., Stricker, N., Liebrecht, C., Hiller, A., Ziegler, M., Lanza, G., 2016. Preparation of Papers for IFAC Conferences & Symposia: Migration Planning for Global Production Networks using Markovian Decision Processes. IFAC-PapersOnLine 49 (12), 35–40.

[24] Rittstieg, M. Einflussfaktoren der Leistungsfähigkeit von Produktionsstandorten in globalen Produktionsnetzwerken. Dissertation, 1. Auflage ed., 266 pp.

[25] Treber, S., 2020. Transparenzsteigerung in Produktionsnetzwerken: Verbesserung des Störungsmanagements durch verstärkten Informationsaustausch, Karlsruhe, 263 pp.

[26] Bramer, W.M., Jonge, G.B. de, Rethlefsen, M.L., Mast, F., Kleijnen, J., 2018. A systematic approach to searching: an efficient and complete method to develop literature searches. Journal of the Medical Library Association : JMLA 106 (4), 531–541.

[27] Stamm, H., Schwarb, T.M., 1995. Metaanalyse. Eine Einführung. Zeitschrift für Personalforschung / German Journal of Research in Human Resource Management 1 (9).

[28] Mongeon, P., Paul-Hus, A., 2016. The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics 106 (1), 213–228.

[29] Chau, K.W., Wu, C.L., 2010. A hybrid model coupled with singular spectrum analysis for daily rainfall prediction. Journal of Hydroinformatics 12 (4), 458–473.

[30] Montgomery, Douglas C. & Peck, Elizabeth A. & Vining, G. Geoffrey. Introduction to Linear Regression Analysis (Wiley Series in Probability and Statistics).

[31] Hui, E.G.M., 2019. Learn R for Applied Statistics. Apress, Berkeley, CA, 254 pp.

[32] Awad, M., Khanna, R., 2015. Efficient learning machines: Theories, concepts, and applications for engineers and system designers. Apress Springer Science+Business Media, New York, NY, 248 pp.

[33] Ruta, D., Gabrys, B., 2007. Neural Network Ensembles for Time Series Prediction. International Joint Conference on Neural Networks.

[34] Martínez-Álvarez, F., Troncoso, A., Morales-Esteban, A., Riquelme, J.C., 2011. Computational Intelligence Techniques for Predicting Earthquakes: Martinez-Alvarez, F; Troncoso; A; Morales-Esteban, A; Riquelme, J.C; 6679, 287–294.

[35] Quinlan, J.R., 1992. Learning with Continuous Classes. AI, 343–348.

[36] Wang, H.R., Wang, C., Lin, X., Kang, J., 2014. An improved ARIMA model for precipitation simulations. Nonlin. Processes Geophys. 21 (6), 1159–1168.

[37] Gui, J., Moore, J.H., Kelsey, K.T., Marsit, C.J., Karagas, M.R., Andrew, A.S., 2011. A novel survival multifactor dimensionality reduction method for detecting gene-gene interactions with application to bladder cancer prognosis. Human genetics 129 (1), 101–110.

[38] Kim, D., Li, R., Dudek, S.M., Ritchie, M.D., 2015. Predicting censored survival data based on the interactions between meta-dimensional omics data in breast cancer. Journal of biomedical informatics 56, 220–228.

[39] Ranstam, J., Cook, J.A., 2018. LASSO regression. British Journal of Surgery 105 (10), 1348.

[40] Ma, S., Fildes, R., Huang, T., 2016. Demand forecasting with high dimensional data: The case of SKU retail sales forecasting with intra- and inter-category promotional information. European Journal of Operational Research 249 (1), 245–257.

[41] LaValley, M.P., 2008. Logistic regression. Circulation 117 (18), 2395–2399.

[42] Empereur-Mot, C., Guillemain, H., Latouche, A., Zagury, J.-F., Viallon, V., Montes, M., 2015. Predictiveness curves in virtual screening. Journal of cheminformatics 7, 52.

[43] Qing Shi, J., Choi, T., 2011. Gaussian process regression analysis for functional data. Chapman & Hall, Boca Raton, 196 pp.

[44] Yan, J., Li, K., Bai, E., Yang, Z., Foley, A., 2016. Time series wind power forecasting based on variant Gaussian Process and TLBO. Neurocomputing 189, 135–144.

[45] Frank, E., Trigg, L., Holmes, G., Witten, I.H., 2000. Technical Note: Naive Bayes for Regression. Machine Learning 41 (1), 5–25.

[46] Fard, M.J., Wang, P., Chawla, S., Reddy, C.K., 2016. A Bayesian Perspective on Early Stage Event Prediction in Longitudinal Data. IEEE Trans. Knowl. Data Eng. 28 (12), 3126–3139.

[47] Yokuma, J., Armstrong, J., 1995. Beyond accuracy: Comparison of criteria used to select forecasting methods. International Journal of Forecasting 11 (4), 591–597.

[48] Armstrong, J.S., 2009. Selecting Forecasting Methods. SSRN Journal.

[49] Domingos, P., 2012. A few useful things to know about machine learning. Commun. ACM 55 (10), 78–87.

Biography

Prof. Dr.-Ing. Dipl.-Wirt. Ing. 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. He is the founder of Schuh & Co. Group and the e.GO Mobile AG. 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. He is Chief Engineer and Head of the Production Management Department at the Laboratory for Machine Tools and Production Engineering WZL at RWTH Aachen University.

Julian Ays (*1990) studied Business Administration and Engineering specializing in Mechanical Engineering at the RWTH Aachen University 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.

Miha Podbreznik (*1995) 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 Global Production group in the Production Management department.

2nd Conference on Production Systems and Logistics

Framework for describing functions of digital technologies in manual assembly systems

Stefan Leineweber[1], Martin Sudhoff[1], Christopher Prinz[1], Bernd Kuhlenkötter[1]
[1]*Lehrstuhl für Produktionssysteme, Ruhr-Universität Bochum, Germany*

Abstract

The integration of digital technologies to promote digitization is essential for the competitiveness of manufacturing enterprises. As part of the production process, the manual assembly is subject of this digitizing change process. According to recent studies the integration of digital technologies is primarily limited to scenarios in larger enterprises or prototype applications. For companies with a manual assembly segment at a lower level of digitization these scenarios, adequately aggregated, potentially contain important insights related to the choice of appropriate digital technologies. An investigation of these scenarios, for example in the form of a qualitative study, therefore seems appropriate. In advance of such an investigation, however, the scope of the investigation, the manual assembly area, must be modeled in order to be able to depict specific situations in the scenarios under investigation. Secondly, the spectrum of digital technologies currently used in manual assembly must be identified and described. The latter is the topic of this paper. The authors of this paper aims to create a framework for describing and classifying digital technologies in manual assembly on the basis of their functions.

Keywords

Digitization; Digital Technologies; Manual Assembly; Framework; Transformation

1. Introduction

Digitization and the associated use of digital technologies are not very advanced yet in the area of manual assembly [1]. In particular, the selection of matching digital technologies (related to the company's goals, the personnel and organizational situation in the assembly area, etc.) represents a challenge for many enterprises [2], [3]. Recommendations for supporting the selection of digital technologies, generated by concepts, models and consulting services, are available in abundance. However, these come with certain disadvantages. The existing consulting services invole high costs [4], are only available regionally or are limited in terms of their scope (e.g. assistance systems, data acquisition) [5], [6]. Self-assessment approaches focus either on a larger domain (enterprise, production area) or a smaller domain (specific work station) and therefore do not allow any precise conclusions to be drawn about the selection of digital technologies in manual assembly [5]. The challenge of providing specific recommendations for action by self-assessment approaches with regard to the identification of appropriate technologies is partly due to the diversity of digital technologies. In addition, there are a large number of factors that characterize the individual situation of a company operating a manual assembly system. [7] However, there are already a number of documented (e. g. [8], [9]) and also non documented best practice cases. An examination of these cases can provide insights into how the enterprises conducted the technology selection process and what factors they considered in the process. These findings can be used to develop an approach that enables other enterprises with manual assembly systems to select the right digital technology for their use case. In this context, a form of explorative research seems to be appropriate [10]. The reason for this is the not too high state of knowledge (although not a low one) regarding the selection of digital technologies for manual assembly. Also, an investigation of this type would strive to gain in-depth understanding of the selection process of digital technologies, which also supports an exploratory investigation. In addition, according to preliminary research, the number of available cases is small. In sum, a qualitative investigation in the form of a case study is considered an appropriate form of investigation. In the apron of this investigation apart from the definition of the general boundary conditions the considered scope of investigation in form of the manual assembly system must be described in a structured way. Another essential aspect is an identification and description of the digital technologies that are potentially applicable in this investigation scope and their functions and properties.

In this paper, as a first step for this qualitative investigation, the digital technologies potentially applicable in the manual assembly area and their functions and properties are described in the form of a framework. For this purpose, a clear understanding of the subject areas and terms will be presented first. This involves digitization (2.1) and digital technologies (2.2) as well as the area of manual assembly (3) as the scope under consideration. This is followed by a summary of the relevant findings from the previous sections and a description of the boundary conditions that apply in further explanations (4). This is followed by an analysis of the functions used by digital technologies in manual assembly (5), taking into account the defined boundary conditions. Finally, the framework developed is presented in the same section.

2. Understanding of digitization and digital technologies

2.1 Digitization

First of all, it can be stated that the term digitization is not clearly defined [11]. However, a distinction can generally be made between two common forms of interpretation of this term. On the one hand, the original understanding in the form of the transformation of analog data into digital data [12]. On the other hand, digitization is understood as a transformation process or trend related to a specific system [13]. A system in this context can represent an entire society, a business network, an enterprise, an enterprise department or just a work system. Digital technologies are an essential part of this transformation process [13], [14], as well as the the transformation from analog data to digital data.The intention of the digitization process or the associated use of digital technologies can be, firstly, the automation of processes in the system to be transformed, or, secondly, a more human-centric approach in which humans are supported in the execution of processes through the use of digital technologies [15], [16]. Since this transformation process is individualized depending on the characteristics of its reference system [3], [17] there is no generally applicable description of a procedure or universal representation of a digitized enterprise or department in the literature [16].

2.2 Digital technologies

The term “digital technology” is a comprehensive term for various technologies. The term is further associated with technologies from the fields of information and communication technologies, identification technologies and automation [18]. However, there is no uniform understanding of digital technologies that describes which specific technologies fall under this term. Commonly named specific technologies in this context include cloud computing, big data, 3D printing, and sensors [14], [19], [20]. In addition to these technologies, a large number of other technologies are classified as digital technologies. They are used to achieve objectives associated with digitization, such as increasing transparency [21], improving the flow of information and increasing flexibility, but also to pursue networking or rationalization purposes within a system [13]. Definite manifestations of these digital technologies are technical systems, often referred to as digitization solutions or digitization applications. In order to create a uniform understanding of digital technologies, [18] have developed a framework, which describes digital technologies on the basis of their general functions in the context of the lean production system. These functions are the storage of data (Storing), the processing of data (Processing), the gathering of data (Gathering), the provision of data and information (Providing), and the transmission of data and information (Communicating) [18].

3. The department of manual assembly

The term "manual assembly system" is used here as the reference system for the subject of manual assembly. The assembly system is a subsystem of production in which objects are assembled into products or assemblies [22]. It is a socio-technical system - a system consisting of interdependent technological, organizational and personnel subsystems [23]. The assembly system itself is composed of assembly stations, which are socio-technical systems likewise [22], [24]. The organizational form of the assembly system is determined by the movement of the assembly object, the movement of the assembly station and the degree of work distribution [25],[26], [27]. Within the assembly stations, the assembly task on an assembly object is performed by human operators and technical systems as part of an assembly process (Figure 1) [28].

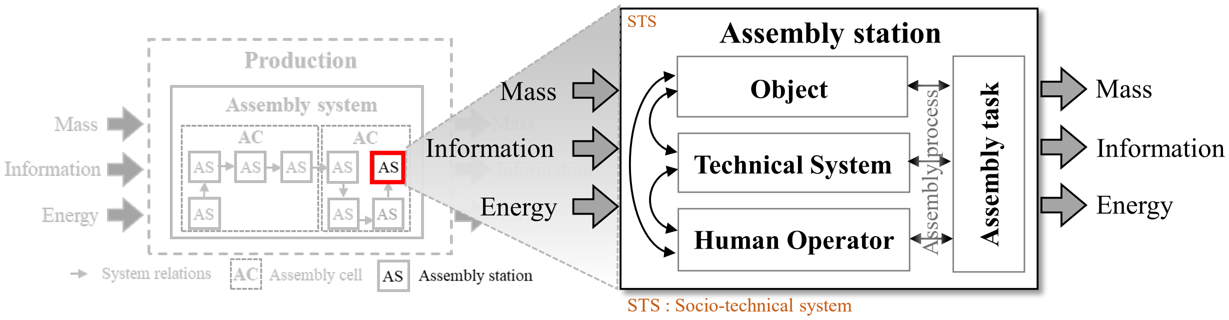


Figure 1: Illustration of an assembly station in relation to an assembly system

Within this assembly process, the sub-functions of the assembly function (joining, handling, controlling, adjusting and special operations) are performed [25], [26]. Other functions that support the assembly process are provided for example by the material supply system and the information system. The material supply system is associated with the functions of storage, commissioning, and transportation [22], [25] whereas the information system supports the function of assembly with the functions of acquisition, provision, communication and processing, as well as storage of data [29].

4. Conclusions and definition of boundary conditions

Section 2.2 explained that there is no general consensus on which explicit technologies fall under the term "digital technology". An approach was outlined in [18] which describes digital technologies on the basis of their functions. This approach and the functions described therein (gathering, providing, storing, processing and communicating of data and information) are adopted in the following to describe the digital technologies relevant for manual assembly systems. An advantage of this approach is the possibility of grouping several technologies under one function [30], [31], [32]. Furthermore, this ensures a long-term validity of the description [18]. It can be assumed that digital technologies will change in their explicit form. However, the functions of technologies are presumably subject to a lower frequency of change.

In order to restrict the scope of the investigation and to reduce complexity, only the concept of a "manual assembly system" will be considered in the following. This includes all relevant inner elements, tasks and processes as well as the interfaces to other relevant systems. The information system and the material supply system are therefore not considered directly in the scope of the investigation.

5. Development and description of the framework

In this section, an analysis of the relevant and currently used individual functions of digital technologies in manual assembly systems is carried out. This forms the basis for the intended development of the framework. The general functions of digital technologies have already been described in sections 2 and 4. In the following, digital technologies are identified, specified and described with regard to their occurrence and their characteristics in the manual assembly system in the context of the set boundary conditions (Figure 2).

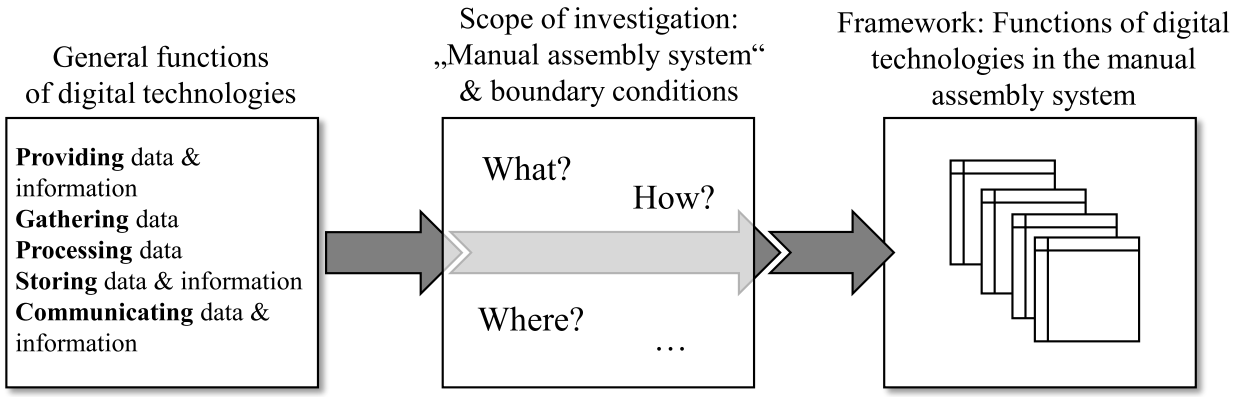


Figure 2: Description of the approach to the development of the framework

5.1 Functions of digital technologies in manual assembly systems

The functions to be identified and described occur in existing practical applications and can therefore be consideres as “known functions”. Accordingly, a detailed description of the functions of relevant digital technologies can and should be provided, which is also in the sense of the informative value of the subsequent qualitative investigation. A rather abstract or iconic representation of functions as in the context of a universal technology selection process (e. g. [31] and [33]) is not pursued.

Identifying and describing the functions of digital technologies requires a structured **procedure**. Therefore, an orientation to the attributes for characterizing work functions according to [34] has been carried out. According to this approach, a work function can be characterized by type (“What?”, “How?”, & “Who?”), duration (“How long?”), occurrence behavior (“When?”) and place of occurrence (“Where?”). However, a work function is always performed by a human. Since in the context of this paper functions of technologies and not of humans are discussed, adjustments have been made to these characteristics. Accordingly, only the following attributes are relevant for characterizing functions of digital technologies: The **type of function (“What?” & “How?”)** and the **place of occurrence of the function (“Where?”)**.

Literature research was mainly conducted to identify the relevant functions. Common databases and platforms (e. g. Science Direct, Research Gate, Springer) were screened using keywords such as "digitalization "+"manual assembly" and "Industry 4.0"+"manual assembly". In addition, current and relevant research projects and related publications were scanned on the websites of project managers and funding organizations. Furthermore, an initial review of relevant and documented projects from consulting institutions such as the competence centers active in Germany was carried out. Further information was obtained from observations of the manual assembly line operated in the learning factory of the Chair of Production Systems in Bochum.

In the previous section, the **boundary condition** was formulated that only the manual assembly system and its interfaces to other systems would be considered in this investigation. This results in an initial reduction of the relevant general functions of digital technologies (Figure 3).

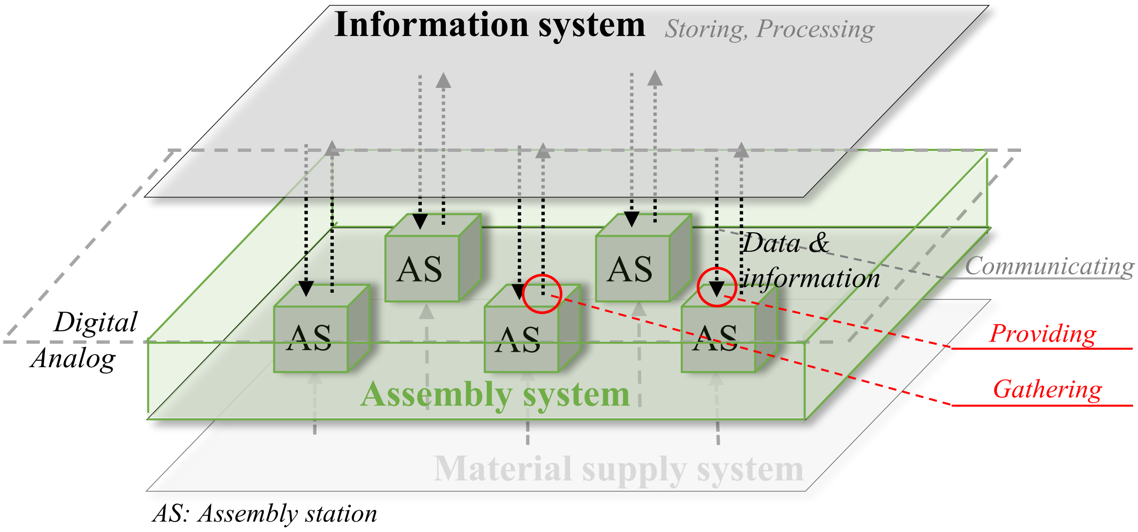


Figure 3: Locating the general functions of digital technologies in the scope of the investigation (marked red)

The functions of processing, storing and communicating data and information are assigned to the information system [29] and are therefore not considered further in the following functional description. Gathering data and providing data and information remain relevant for further discussion. These two functions will now be identified and specified according to type and location of occurende as part of the procedure already described.

The question of **"What?"** in the context of characterizing the **type of function** has already been answered in general terms with "data" in the case of the "gathering" function. In order to describe the functions in the context of manual assembly as precisely as possible, it is necessary to specify in more detail below what data is actually involved. Since the number of generally gathered data in manual assembly systems is large and the manifestations of these data are manifold [3], a classification of the relevant data is carried out first. In this context, some approaches already exist (e.g. [6], [35] [36] and [37]). Based on these approaches, the identification and classification is performed. The classification is done with respect to the reference element of the data. A distinction is made between order-related, assembly object-related, process-related, human-related and logistics-related data. The following table (Table 1) shows some concrete examples of these data classes.

Table 1: Classification and specific data in manual assembly systems in context of the function „Gathering“

Classification	Specific data
order-related	order status, quantities, times, …
assembly object-related	weight, dimensions, image, …
process-related	malfunctions, worker experience, reject, proportion of good parts, unloading position, unloading sequence, joining force, …
human-related	satisfaction, ergonomic strain, attendance, …
logistic-related	stock at the assembly station, weight of the carrier, number of unloading processes, number of insertion processes, missing parts, ...

Furthermore, the **type of function ("How?")** is specified. Existing approaches from the field of digital assistance systems as well as from the areas of sensor technology and production data acquisition form the basis for own explanations. Examples are [28], [35], [38], [39], [40] and [41]. Taking these approaches into account, a distinction is made here between optical data acquisition (e.g. by cameras or light barriers), acoustic data acquisition (e.g. by microphone), mechanical data acquisition (e.g. by scales, vibration sensors) and electromagnetic data acquisition (e.g. via interfaces). Furthermore, in relation to the type of function and the question of "How?" the degree of automation of the acquisition is differentiated. Here, a differentiation is made between manual acquisition and automated acquisition [35], [39], [42].

Finally, the **potential occurrence locations ("Where?")** in the manual assembly system were identified for the "**Gathering**" function. These are mainly one or several assembly stations. However, in addition to this decentralized approach, data can also be captured from a central location in the assembly system (e. g., central terminal for capturing order data). [43]

After the discussion of the "Gathering" function, the identification and description of the function of **providing data and information** in manual assembly systems ("**What?**") follows. Here, again, the first step is to identify which specific data and information are available according to literature and current application scenarios. These were also assigned to specific classes. In analogy to the "Gathering" function, the data and information are also classified in the case of "Provision" function with regard to their reference element in the manual assembly system. Reference sources for the own explanations are for example [28], [29], [40] and [44]. Data and information are assigned to the following classes: Order-related, assembly object-related, assembly task-related, process-related and logistics-related. The following table (Table 2) lists these classes with specific examples of data and information.

Table 2: Classification and specific data and information in manual assembly systems in context of the function „Providing"

Classification	Specific data and information
order-related	order number, quantities; worklist, planned order start, planned order completion, key performance indicators (e. g. current productivity), …
assembly object-related	bill of material, construction drawing (2D, 3D), circuit diagramm, images, …
task-related	assembly instruction, work plan, additional information (e. g. worker experience), test plan, assembly part position, force to be applied, …
process-related	abrasion, malfunctions, …
logistic-related	stock, object position, …

Also, an identification of the **type of provision ("How?")** common in manual assembly systems is performed. A distinction is made here between visual, acoustic and haptic provision [38], [40] [45]. Furthermore, following the approaches [40], [46] and [47], the type of provision is differentiated into manual or automatic provision.

Once again, individual or several assembly stations have been identified as the **place where data and information are provided ("Where?")**. However, the provision can, just like the gathering, also take place at a central location in the assembly system. [8], [48]

5.2 Presentation of the framework

The created framework shows potentially applied and relevant functions that can be performed by digital technologies in manual assembly systems according to the state of the art. The framework also takes into account that the design of the functions "provision" and "acquisition" is specific depending on the information and data under consideration. For example, not all information is suitable for acoustic provision. Reasons for this may be differences in the complexity of the information. In the following table (Table 3) an overview of the created framework is presented. The entire filled-in framework cannot be illustrated in its entirety within the bounds of the paper due to its size and the probably resulting lack of readability. The framework is a tool that will be used in the subsequent qualitative investigation. In the course of the investigation of individual cases, functions of digital technologies can be classified in this framework.

Furthermore, this framework can be integrated into a concept to support the selection of digital technologies for manual assembly.

Table 3: Overview illustration of the framework

Function: Gathering				
Function type				**Function location**
"What?"		***"How?"***		***"Where?"***
Data		**Form**	**Level of automation**	
Classification	***Specific***			
order-related	order status	optical, acoustical	manual, automated	central, decentralized: at 1-n assembly stations
	…	…	…	…
assembly object-related	weight	mechanical, electromagnetical	manual, automated	decentralized: at one assembly station
…	…	…	…	…

Function: Providing				
Function type				**Function location**
"What?"		***"How?"***		***"Where?"***
Data & information		**Form**	**Level of automation**	
Classification	***Specific***			
order-related	work list	visual, acoustical	manual, automated	central, decentralized: at 1-n assembly stations
	…	…	…	…
task-related	assembly instruction	visual	manual, automated	decentralized: at 1-n assembly stations
…	…	…	…	…

6. Conclusion and outlook

In the context of this paper, a framework for describing the functions of digital technologies currently used in manual assembly systems was carried out in preparation for conducting a qualitative investigation. Further adaptations and extensions of the framework are possible in order to be able to carry out this qualitative investigation more specifically.

Following on from this, a model for mapping the situation in manual assembly needs to be set up, in which potential factors influencing the selection of digital technologies are outlined. In the context of the actual qualitative investigation, it is then necessary to determine in which constellations of influencing factors which digital technologies are suitable. However, due to the large number of different types of influencing factors, it can be assumed that it will not be possible to examine the interactions of all types in the qualitative investigation. Accordingly, the most important influencing factors must be identified. Furthermore, the already started acquisition of cases will be accomplished and subsequently the actual qualitative investigation will be conducted.

References

[1] Rusch, T.; Riegel, A.; Hueber, M.; Kerber, F.; Merkel, L.; Sochor, R. 2021. Synchronisation von Digitalisierung, Qualitätssicherung und Assistenzsystemen an Arbeitsplätzen mit geringem Automatisierungsgrad. in: Jeske, T.; Lennings, F. (Ed.), Produktivitätsmanagement 4.0. Springer. Berlin Heidelberg. pp. 185–228.

[2] Siedler, C.; Aurich, J. 2018. Digitale Transformation eines Produktionssystems. Entwicklung eines unternehmensindividuellen, modularen Konzeptes zur digitalen Transformation eines Produktionssystems in ein CPPS. In: Zeitung für Wirtschaftlichen Fabrikbetrieb (ZWF) (113), pp. 514–517.

[3] Schuh, G.; Anderl, R.; Dumitrescu, R.; Krüger, A.; Hompel, M. (Ed.) 2020. Industrie 4.0 Maturity Index.UPDATE 2020. Die digitale Transformation von Unternehmen gestalten. Acatech. München.

[4] Tödtmann, C. 2017. Strategieberater sind am teuersten – was Unternehmensberater für welche Expertise kassieren. Wirtschaftswoche. Online verfügbar unter https://blog.wiwo.de/management/2017/01/23/strategieberater-sind-am-teuersten-was-unternehmensberater-fuer-welche-expertise-kassieren/, zuletzt aktualisiert am 23.01.2017, zuletzt geprüft am 26.01.2021.

[5] Fechtelpeter, C.; Heim, Y.; Löffler, T.; Niewöhner, N. 2019. Vorstudie zur Entwicklung einer bedarfs- und nutzergerechten Unterstützung von KMU bei der Einführung und Anwendung von Industrie 4.0. Acatech (Ed.). München.

[6] Bischoff, J. (Ed.); Taphorn, C.; Hegemanns, T.; Prasse, C.; Braun, S.; Guth, M. 2015. Studie - Erschließung der Potenziale der Anwendung von "Industrie 4.0" im Mittelstand. Studie im Auftrag des Bundesministeriums für Wirtschaft und Energie. Mühlheim an der Ruhr.

[7] Ullrich, A.; Plank, M.; Weinert, N. 2017. Handlungsempfehlungen. In: Ullrich, A. (Ed.): Metamorphose zur intelligenten und vernetzten Fabrik. Ergebnisse des Verbundforschungsprojekts MetamoFAB. Springer Vieweg. Berlin, Heidelberg. S. 211–216.

[8] Digital in NRW 2018. Digitalisierung im Schaltschrankbau. Projekt mit der Schaltanlagenbau GmbH H. Westermann. Online available at: https://www.digital-in-nrw.de/de/erfolgsgeschichten/qualitaets-und-prozessmanagement/digitalisierung-im-schaltschrankbau, last updated on 08.04.2020, last reviewed on 08.04.2020.

[9] Digital vernetzte Montagelinie - Digital in NRW 2020. Online available at https://www.digital-in-nrw.de/de/erfolgsgeschichten/produktionsplanung-und-steuerung/digital-vernetzte-montagelinie, last updated on 08.04.2020, last reviewed on 08.04.2020.

[10] Albers, S.; Klapper, D.; Konradt, U.; Walter, A.; Wolf, J. 2007. Methodik der empirischen Forschung. 2., überarbeitete und erweiterte Auflage.: Springer Fachmedien. Wiesbaden.

[11] Anderl, R. 2015. Leitfaden Industrie 4.0. Orientierungshilfe zur Einführung in den Mittelstand. VDMA-Verlag. Frankfurt am Main.

[12] Bitkom 2016. Industrie 4.0 - Status und Perspektiven. Bitkom (Ed.). Berlin.

[13] Hertwig, M. 2020. Digitalisierung der Wissensgesellschaft. Paradoxien des technologischen Wandels im Zeitalter von Crowdsourcing und Industrie 4.0. In: Kouli, Y.; Pawlowsky P.; Hertwig, M. (Ed.): Wissensökonomie und Digitalisierung. Springer Fachmedien Wiesbaden, pp. 129–156.

[14] Hanschke, I. 2018. Digitalisierung und Industrie 4.0 - einfach und effektiv. Systematisch & lean die Digitale Transformation meistern. Hanser. München.

[15] Klimant, P. 2020. Industrie 4.0 in der industriellen Praxis. In: Kouli, Y.; Pawlowsky P.; Hertwig, M. (Ed.): Wissensökonomie und Digitalisierung. Springer Fachmedien. Wiesbaden. pp. 157–166.

[16] Klotz, U. 2018. Zukunft der Arbeit. In: Barton, T.; Müller, C.; Seel, C. (Ed.): Digitalisierung in Unternehmen. Wiesbaden: Springer Fachmedien Wiesbaden, pp. 12–25.

[17] Westkämper, E.; Löffler, C. 2016. Strategien der Produktion. Springer Berlin Heidelberg.

[18] Hoellthaler, G.; Meister, F.; Braunreuther, S.; Reinhart, G. 2020. Function framework for describing digital technologies in the context of lean production. In: Procedia CIRP 88, pp. 167–172.

[19] Bundesministerium für Wirtschaft und Energie 2017. Sensor, Tablet, RFID: Digitale Technologien in der Produktion. Ed. v. Bundesministerium für Wirtschaft und Energie. Berlin

[20] Deloitte 2016. Manufacturing 4.0: Meilenstein, Must-Have oder Millionengrab. Warum bei M4.0 die Integration den entscheidenten Unterschied macht.

[21] Herkommer, O.; Kleinbeck, T.; Nitsche, J. 2017. Steigerung der Wettbewerbs-fähigkeit und Menschzentrierung stehen nicht im Widerspruch. In: ZWF 112 (9), pp. 621–625.

[22] Willnecker, U. 2000. Gestaltung und Planung leistungsorientierter manueller Fließmontagen. Dissertation. Technische Universität München, München. Lehrstuhl für Betriebswissenschaften und Montagetechnik.

[23] Hirsch-Kreinsen, H. 2014. Wandel von Produktionsarbeit – Industrie 4.0. In: WSI-Mitteilungen (6), pp. 421–429.

[24] Eversheim, W.; Schuh, G. 1999. Produktion und Management 3. Gestaltung von Produktionssystemen., Heidelberg: Springer. Berlin.

[25] Seliger, G. 2014. Montage und Demontage. In: Grote, K; Feldhusen, J. (Ed.): Dubbel. Springer Berlin Heidelberg, pp. 1461–1467.

[26] Lotter, B.; Wiendahl, H. 2012. Montage in der industriellen Produktion. Springer Berlin Heidelberg.

[27] Schuh, G.; Schmidt, C. 2014. Produktionsmanagement. Springer Berlin Heidelberg.

[28] Schlick, C.; Bruder, Ralph; Luczak, H. 2010. Arbeitswissenschaft. Springer Berlin Heidelberg.

[29] Johansson, P.; Eriksson, P.; Johansson, P.; Malmsköld, L.; Fast-Berglund, Å.; Moestam, L. 2017. Assessment based Information Needs in Manual Assembley. In: 24th International Conference on Production research (ICPR 2017). Posnan, Poland, July 30-August 3, 2017. Lancaster: DEStech Publications. pp. 366–371.

[30] Kröll, M. 2007. Methode zur Technologiebewertung für eine ergebnisorientierte Produktentwicklung. Dissertation. Stuttgart.

[31] Schimpf, S.; Lang-Koetz, C. 2010. Technologiemonitoring. Technologien identifizieren, beobachten und bewerten. Fraunhofer IAO. Stuttgart.

[32] Wolffgramm, H. 2012. Allgemeine Techniklehre. Elemente, Strukturen und Gesetzmäßigkeiten. Einführung in die Denk- und Arbeitsweisen der allgemeinen Techniklehre.

[33] VDI-Richtlinie 2803, Januar 2019: Funktionenanalyse. Beuth Verlag. Düsseldorf.

[34] Schmiedl, N. 2007. Methodik zur prozessorientierten Restrukturierung von Arbeitssystemen. Dissertation. Technische Universität Chemnitz. Institut für Betriebswissenschaften und Fabriksysteme. Chemnitz.

[35] Loos, P. 1999. Grunddatenverwaltung und Betriebsdatenerfassung als Basis der Produktionsplanung und –steuerung. In: H. Corsten und B. Friedl (Hg.): Produktionscontrolling. München: Vahlen, pp. 227–252.

[36] Möller, M. 2018. Betriebsdatenerfassung. Grundlagen, Ziele und Trends auf dem Weg zur Smart Factory.

[37] Aehnelt, M. 2016. Informationsassistenz zur kognitiven Automatisierung manueller Montagearbeitsplätze. Dissertation. Universität Rostock. Rostock.

[38] Hold, P. 2020. Vorgehensmodell zur Planung und Evaluierung digitaler Assistenzsysteme in der Montage. Dissertation. Technische Universität Wien. Wien.

[39] Klapper, J.; Gelec, E.; Pokorni, B.; Hämmerle, M.; Rothenberger, R. 2019. Potenziale digitaler Assistenzsysteme. Aktueller und zukünftiger Einsatz digitaler Assistenzsysteme in produzierenden Unternehmen. Ed. v. Fraunhofer IAO. Stuttgart.

[40] Hinrichsen, S. (Ed.) 2020. Informatorische Gestaltung der Montage mittels Assistenzsystemen. Informatorische Assistenzsysteme in der variantenreichen Montage: Springer Vieweg. Berlin.

[41] Bullinger, H. J.; Lung, M. 1994. Planung der Materialbereitstellung in der Montage. Springer Fachmedien. Wiesbaden.

[42] Acatech 2016. Kompetenzentwicklungsstudie Industrie 4.0. Erste Ergebnisse und Schlussfolgerungen. Acatech. München.

[43] NRW Innovationspartner 2021. Einfach effizienter mit dem Tablet. Online available at: https://www.nrw-innovationspartner.de/aktuelles/erfolgsgeschichten/einfach-effizienter-mit-dem-tablet/, last updated on 26.04.2021, last reviewed on 26.04.2021.

[44] Franken, S.; Prädikow, L.; Vandieken, M. 2019. Fit für Industrie 4.0? Ergebnisse einer empirischen Untersuchung im Rahmen des Forschungsprojektes „Fit für Industrie 4.0". FGW Studie. (Ed.): Hartmut Hirsch-Kreinsen und Anemari Karacic. Digitalisierung von Arbeit, 18. Düsseldorf.

[45] Niehaus, J. 2017. Mobile Assistenzsysteme für Industrie 4.0. Gestaltungsoptionen zwischen Autonomie und Kontrolle. (Ed.): FGW – Forschungsinstitut für gesellschaftliche Weiterentwicklung e.V. Düsseldorf.

[46] Hinrichsen, S.; Riediger, D.; Unrau, A. 2016. Assistance Systems in Manual Assembly. In: Villmer, F. (Ed.): Production engineering and management. Proceedings 6th international conference, September 29 and 30, 2016, Lemgo, Germany. Hochschule Ostwestfalen-Lippe, University of Applied Sciences; Università, Degli Studi di Trieste, pp. 3–14.

[47] Funk, M.; Bächler, A.; Bächler, L.; Korn, O.; Krieger, C.; Heidenreich, T.; Schmidt, A. 2015. Comparing projected in-situ feedback at the manual assembly workplace with impaired workers. In: Fillia Makedon (Ed.): Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '15. Corfu, Greece, 01.07.2015 - 03.07.2015. ACM Press. New York.

[48] Digital in NRW - Kompetenz für den Mittelstand 2021. Familienunternehmen setzt auf datenbasierte Verbesserung der Produktions-transparenz. DAISY: Digital in NRW und Horstmann GmbH entwickeln neues Tool. Online available at: https://www.digital-in-nrw.de/de/hauptnavigation/erfolgsgeschichten/produktionsplanung-und-steuerung/datenbasierte-verbesserung-der-produktionstransparenz, last updated on 21.04.2021, last reviewed on 21.04.2021.

Biography

Stefan Leineweber (*1987) studied mechanical engineering at the Ruhr University Bochum. He is a research associate in the working group Production Management at the Chair of Production Systems at the Ruhr University Bochum. His primary research topic is digitalization.

Martin Sudhoff (*1991) is a research associate at the Chair of Production Systems (LPS) at the Ruhr-University of Bochum since 2017. He earned a bachelor's and master's degree in mechanical engineering at the Ruhr-University of Bochum. His primary research topics are the digitalization and automation of assembly systems.

Christopher Prinz (*1985) studied mechanical engineering at the Ruhr University Bochum. After receiving his doctorate in 2018 on the topic of knowledge management in production, he was named Academic Councilor at the Chair of Production Systems (LPS). As part of the chair management, he is responsible for the strategic development of the chair and the initiation of research projects.

Bernd Kuhlenkötter (*1971) 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. Since 2015 he holds the professorship of the Chair of Production Systems at the Ruhr-Universität Bochum.