RESEARCH PAPER

Predictive maintenance as an internet of things enabled business model: A taxonomy

Jens Passlick[1] · Sonja Dreyer[1] · Daniel Olivotti[1] · Lukas Grützner[1] · Dennis Eilers[2] · Michael H. Breitner[1]

Abstract

Predictive maintenance (PdM) is an important application of the Internet of Things (IoT) discussed in many companies, especially in the manufacturing industry. PdM uses data, usually sensor data, to optimize maintenance activities. We develop a taxonomy to classify PdM business models that enables a comparison and analysis of such models. We use our taxonomy to classify the business models of 113 companies. Based on this classification, we identify six archetypes using cluster analysis and discuss the results. The "hardware development", "analytics provider", and "all-in-one" archetypes are the most frequently represented in the study sample. For cluster analysis, we use a visualization technique that involves an autoencoder. The results of our analysis will help practitioners assess their own business models and those of other companies. Business models can be better differentiated by considering the different levels of IoT architecture, which is also an important implication for further research.

Keywords Taxonomy · Predictive maintenance · Business models · IoT · Cluster analysis

JEL classification L86

Introduction

The introduction of the Internet of Things (IoT), in terms of both theory and practice, is currently the subject of intense discussions (Whitmore et al. 2015). The IoT has enormous potential in both the private and industrial environments (Manyika et al. 2015). The term Industrial Internet of Things (IIoT) is used for such applications. Prior research discusses the characteristics of business models that successfully use the

possibilities offered by IIoT (Herterich et al. 2016). Previous research on the IoT environment shows that understanding the business models of company partners is important for long-term success (Dijkman et al. 2015). Digital business models in general are analyzed in prior research (e.g., Hartmann et al. 2016; Bock and Wiener 2017; Rizk et al. 2018). However, the more general taxonomies used for digital business models include aspects that are not relevant for every company with an IoT or IIoT business model (Bock and Wiener 2017). Particularly in the context of Industry 4.0, in which IIoT is a major component, a more concrete consideration of the changes that have been made to business models is important. In the area of value creation, value offer, and value capture, specific aspects must be considered in Industry 4.0 (Müller and Buliga 2019). Initial research has been conducted on business model archetypes involving Industry 4.0, but current knowledge needs to be deepened to understand the interactions among the different actors involved in value creation networks. Previous research focuses on service-driven business models, but business models in the manufacturing industry are rarely addressed (Müller and Buliga 2019).

In many industrial applications, maintenance is an important factor that is often discussed in terms of cost savings (Khazraei and Deuse 2011). Considering maintenance

✉ Jens Passlick
passlick@iwi.uni-hannover.de

Dennis Eilers
dennis.eilers@deepcorr.com

[1] Information System Institute, Leibniz Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany

[2] DeepCorr GmbH, Göttinger Hof 7, 30453 Hannover, Germany

⚫ Springer

services, a study has shown that the potentially most valuable action is predictive maintenance (PdM) (Holgado and Macchi 2014). PdM applications represent one way of using the IIoT to reduce costs. PdM uses data, especially sensor data obtained from IoT devices, to optimize maintenance activities. Often, this process also includes "condition monitoring" (Khazraei and Deuse 2011; Borgi et al. 2017).

The aim of PdM is not to carry out maintenance unnecessarily early or too late, and involves being able to make forecasts about the further deterioration of, e.g., a machine. In particular, unscheduled deterioration can be detected so the operator can act proactively. The various aspects of PdM are considered in the following definition: PdM is "condition-based maintenance carried out following a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item" (BSI (British Standards Institution) 2010, p. 12). In the next section, we further explain the aspects of PdM.

The consulting firm PricewaterhouseCoopers, in cooperation with Mainnovation, concluded that out of 280 surveyed companies in Germany, Belgium and the Netherlands, 132 companies might want to implement a PdM solution, and 52 companies are already working on such an implementation (PricewaterhouseCoopers 2017). These numbers show the relevance of PdM for companies. The relevance of PdM is also increasing in the scientific field (Daily and Peterson 2017). However, it is difficult for companies to understand the market of PdM providers and offerings. Which providers are on the market and what do they offer? Research specifically considering the characteristics of digital business models for maintenance services is lacking. In a scientific discussion of PdM business models, it is important to consider their different forms to better understand how they work in practice. An investigation of the extent that general taxonomies for digital or IoT business models can be applied to a specific IoT application would contribute to current research. By using a taxonomy for a specific application, companies can better identify other companies using their own business models on the market as well as potential growth opportunities. Therefore, we develop the following research question:

Which elements of PdM business models are important and which characteristics are interrelated in models that exist on the market?

The article proceeds as follows: First, we describe and define PdM and discuss the related literature. Next, we develop a taxonomy for PdM business models using a procedure proposed by Nickerson et al. (2013). Then, we use the final taxonomy to classify the business models of 113 companies, conduct a cluster analysis and build archetypes that represent typical PdM business models. Finally, we discuss our results, outline the implications and limitations, and suggest further research.

Predictive maintenance and the related literature

Comprehensive insight into the current condition of a component or machine is necessary for PdM (Sipos et al. 2014). In a broader sense, data are key for PdM (Borgi et al. 2017). Usually, a central server is used to collect, transmit and process data (Wang et al. 2017). Data collection must take place often in (near) real time, which is why control tools that are capable of collecting data automatically from several components and systems are useful (Aivaliotis et al. 2017). Data must not only be recorded but also processed and analyzed (Cachada et al. 2018). Appropriate tools not only process sensor data but also take the maintenance history, operational data, design and application into account (Darwanto et al. 2012).

Monitoring and determining the current state of equipment is the first step of PdM (Hui et al. 2008). The aim is to detect the beginning of degradation as early as possible (Borgi et al. 2017; Khazraei and Deuse 2011). To achieve this, for PdM to be reliable, all information must be recorded. Sensors can be used to record condition-related data (Sipos et al. 2014). Indicators must be identified, measured and modeled so that the corresponding activities can be derived (Groba et al. 2007). Vibration analysis, thermal images (Barbera et al. 1996), trend analysis and simulation (Aivaliotis et al. 2017) are examples of techniques that are used. The identified elements are summarized in general IoT architectures (e.g., Chen 2013; Turber et al. 2014).

In the present article, we understand PdM as the most comprehensive form of maintenance that includes condition-based maintenance and additional types of maintenance that are enabled by data analysis. Condition-based maintenance is different from PdM because decisions are based only on the current condition of the focal object. In contrast, prediction tools and methods are used in PdM (Susto et al. 2012). In a previous article, condition-based maintenance and PdM are equated (Last et al. 2010). Khazraei and Deuse (2011) state that avoidance-based, condition-based, and detective-based maintenance are tactics used for PdM. In the following, we also take this approach. Thus, PdM is condition monitoring with "a forecast derived from repeated analysis or known characteristics and evaluation of the significant parameters of the degradation of the item" (BSI (British Standards Institution) 2010, p. 12).

The use of PdM has various advantages. PdM minimizes system downtime, leading to a reduction in production losses (Baidya and Ghosh 2015; Spendla et al. 2017; Zoll et al. 2018). In contrast to regularly performed maintenance activities, PdM takes the current condition of the system into account (Chu et al. 1998). This consideration leads to a reduction in maintenance activities (Last et al. 2010; Susto et al. 2013). Maintenance activities are only performed when required as long as the system is still running in its intended way (Mattes et al. 2012). PdM reduces the probability of extensive failures (Darwanto et al.

2012). From an economic perspective, costs are reduced because the maintenance activities are less precise, and there is less likelihood of default (Wang et al. 2009). Additionally, the customer experience is enhanced, and customer loyalty is strengthened (Gerloff and Cleophas 2017).

We conducted a structured literature review using the guidelines proposed by Webster and Watson (2002) and vom Brocke et al. (2015) to identify existing taxonomies in the field of PdM. We searched the following databases: the Digital Library of the Association for Computing Machinery (ACM), The Association for Information Systems e-Library (AISeL), Emerald Insight, the Institute of Electrical and Electronics Engineers (IEEE) Xplore digital library, InformsOnline, JSTOR, ScienceDirect, and SpringerLink. The search terms used were, namely, "*taxonomy*", "*maintenance*", and "*predictive maintenance*". We also searched using the strings "*taxonomy condition monitoring*", "*condition monitoring business models*", and "*predictive maintenance business model*". We thereby identified five articles. Forward and backward searches led to four additional results. To search for particularly relevant articles, Google Scholar was also used. In addition to the abovementioned articles, three other articles were identified.

Among the relevant studies, some articles discuss the different components of PdM business models. However, these articles do not explicitly deal with embedding these aspects in a business model. Khazraei and Deuse (2011) develop a classification for different maintenance types with the objective of simplifying technical communication. Rizk et al. (2018) identify the characteristics of data-driven digital services. These scholars consider cross-organizational contexts and highlight the complexity of digital services in their article. A special form of digital service is analyzed by Hunke et al. (2019). These scholars develop a taxonomy of analytics-based services. In addition to dimensions such as "data generator", "data target", and "data origin", they observe that analytical services can be divided into descriptive, diagnostic, predictive, and prescriptive analyses. Herterich et al. (2016) describe a taxonomy that classifies industrial service systems enabled by digital product elements. The taxonomy allows the identification of properties to be changed to exploit digital potentials. The taxonomy does not specifically focus on business-to-business (B2B) or business-to-customer (B2C) applications.

While previous articles do not yet address IoT or data-driven business models, Bock and Wiener (2017) conduct a general analysis of digital business models with their taxonomy, where the important dimensions are the value promise and pricing strategy. Data-driven business models are also examined by Engelbrecht et al. (2016). Based primarily on the data source, these scholars identify eight different categories of business models without focusing on a specific industry. However, they focus on startups, and established companies are not a main consideration. These scholars argue that developing a new business model is simplified through the taxonomy by looking at the respective category of business models (Engelbrecht et al. 2016). Hartmann et al. (2016) develop a taxonomy for startups using data-driven business models and categorize data sources, key activities, target customers, revenue models and cost advantages. This taxonomy forms the basis for a framework that can be used to create new business models in the field of big data. Täuscher and Laudien (2018) examine platform business models by looking at the key values, i.e., price-cost-efficiency, emotional value and social value, among other things. They consider both B2C and B2B applications but do not focus mainly on the IoT. Specific IoT platforms are analyzed by Hodapp et al. (2019). They develop a taxonomy for business models using IoT platforms.

The articles discussed so far either do not deal with business models, do not explicitly consider the IoT or have a specific focus on IoT platform operators. In contrast, Paukstadt et al. (2019) develop a taxonomy of smart services, considering both B2C and B2B smart services. However, the examined data set contains more B2C applications than B2B applications. Smart services are classified based on the concept, delivery and monetization. Weking et al. (2018) develop a taxonomy for Industry 4.0 business model innovations, aiming to describe the transition from traditional business models to business models using Industry 4.0. These scholars consider manufacturing companies, whose customers can be both end consumers and other companies. Müller and Buliga (2019) investigate data-driven business models in the context of Industry 4.0. They provide an initial overview of the archetypes of data-driven business models used in the B2B environment, and the three archetypes are differentiated only by value creation, value offer, and value capture. These scholars suggest that a deeper analysis "of data-driven business model innovation in B2B contexts" is needed (Müller and Buliga 2019, p. 6).

In summary, there is research on IIoT and Industry 4.0 business models, but there are only a few articles, and they are not very detailed. In particular, there is less research on the B2B context based on empirical data. Different maintenance strategies are already classified but these classifications are not as detailed as could be if a taxonomy is used. A taxonomy can show how diverse PdM offers are by considering existing offers from companies all over the world. Although many articles deal with PdM and digital business models, a unifying, comprehensive taxonomy is not yet available. Taxonomies enable a better understanding of a research field and provide a foundation for theory building (Szopinski et al. 2019). Furthermore, taxonomies help identify and explain differences and similarities (Nickerson et al. 2017). For the case of PdM, the individual elements of the business models can be identified, and their relations can be examined (Glass and Vessey 1995).

Taxonomy development

Procedure

In the information systems (IS) field, classification systems, including taxonomies, are theoretical artifacts for describing and analyzing the characteristics of objects and their relationships (Gregor 2006). Taxonomies are widely applied in IS research, but until 2013, a structured methodology was often missing, and taxonomies were developed using an intuitive approach (Nickerson et al. 2013). Nickerson et al. (2013) published a structured and accepted process for the development of taxonomies, which we followed. Their process is based on Bailey's (1984) three-level indicator model and Hevner et al. (2004)'s design-science research guidelines. The term "taxonomy" is defined as "a set of it n dimensions D_i (i=1, ..., n) each consisting of k_i (k_i>2) mutually exclusive and collectively exhaustive characteristics [...]." (Nickerson et al. 2013, p. 340). Therefore, the taxonomy development takes place in several iterations (Nickerson et al. 2013). In each iteration a different approach is conceivable. Either the taxonomy is based on concepts (conceptual-to-empirical), usually involving existing models, or empirical data (empirical-to-conceptual). During the development of the taxonomy, the focus is usually on a certain area of interest, which is determined as a meta-characteristic at the beginning of the process. This meta-characteristic is a superordinate and abstract description of the area on which the taxonomy focuses, and serves as the basis for the dimensions and characteristics used in the taxonomy. Fig. 1 shows this procedure.

In our case, the meta-characteristic is used to define the elements of PdM business models. Based on the results of the analysis of scientific literature on business models, the dimensions of the taxonomy are conceptually derived. Next, related characteristics are developed by empirically examining a large number of globally distributed companies active in PdM. According to the definition proposed by Nickerson et al. (2013), the characteristics of a company can be seen as exclusive. Exclusive means that in each dimension, exactly one characteristic is assigned to each company. After each iteration, a decision is made based on various end conditions to determine whether a further iteration is necessary. The end conditions used are adapted from Nickerson et al. (2013), see Appendix Table 5. The following section describes the steps performed in each iteration.

Iterations

For the first iteration, a conceptual-to-empirical approach was used. Nickerson et al. (2013) advise starting with the conceptual-to-empirical approach "if little data are available but the researcher has significant understanding of the domain" (p. 345). Since there is little known about the structure of PdM business models, we have oriented ourselves to the literature on general business models. The analysis of this literature involved a review of existing knowledge and the identification of key terms relevant to our taxonomy. The findings formed the basis for the first dimensions of the meta-characteristic. We compared the elements of the electronic business models following the process proposed by Afuah and

Fig. 1 Taxonomy development procedure proposed by Nickerson et al. (2013, p. 345)
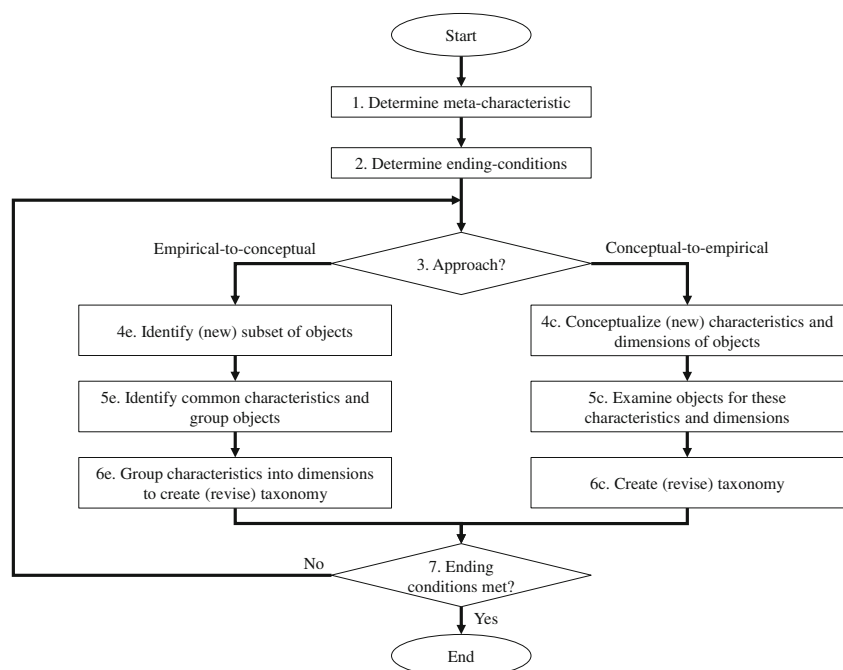
Tucci (2001), Alt and Zimmermann (2001), Brousseau and Pénard (2007), Mahadevan (2000), Osterwalder and Pigneur (2010). The business model canvas tool developed by Osterwalder and Pigneur (2010) summarizes the majority of the elements of business models discussed in the literature. Additionally, the business model canvas is highly regarded in practice. Therefore, we decided to use this as the basis for our taxonomy. Possible dimensions were discarded because many of the PdM business models were similar in that respect (e.g., key resources) or if no relevant information was available (cost structures, key partners). We considered the element "sales channels" to be a useful dimension for differentiating PdM business models. According to Osterwalder and Pigneur (2010), the sales channel represents the methods used to sell a product or service and reach customers. Sales channels include, for example, the use of Internet marketplaces or the use of direct sellers. Furthermore, revenue streams were included to assess the payment models offered to customers. Is paid just once for a complete product or is there a monthly use fee for a specific service? The customer segment dimension was included to describe the customers. The end of the first iteration resulted in the following dimensions: key activities, value proposition, revenue streams, sales channels, and customer segments. Several end conditions of the taxonomy development were not fulfilled after the first iteration (see Appendix Table 5); therefore, another iteration was necessary.

In the second iteration, the empirical-to-conceptual approach was used, and data from real PdM business models were analyzed. We conducted 42 interviews with representatives of various companies at the "Hannover Industrie Messe" 2018, a leading fair for industrial automation and IT technologies. We searched the exhibitor list for companies with the tag "predictive maintenance". We then checked the website of the company to see if their understanding of PdM fit our definition. In the interviews, we discussed different topics on various aspects of the company (the list of topics is provided in the Appendix in the first section). The survey was based on the previously discussed knowledge about business models and the results of the first iteration of the taxonomy development. The interviews lasted between five minutes and 40 min, on average approximately 15 min. In addition, terms including "companies", "predictive maintenance" and "condition monitoring" were searched for on Google to identify PdM companies. During this search, webpages with lists of companies that use PdM business models were found. After the interviews and Google search were conducted, the database included 71 companies. We then used the Crunchbase website (a database created to track startups containing various information about companies) to search for appropriate companies. Further, we used a base account in Crunchbase to download the open data map. In the file, containing information about the companies, the terms "predictive maintenance" and "condition monitoring" were used to search the short descriptions. All companies

identified via these methods were selected according to whether their understanding of PdM aligned with our definition. Using this method, 42 additional companies were identified, which resulted in 113 companies in our entire database (see Appendix Table 7).

Initially, a random sample of ten companies was examined from which suitable characteristics for the dimensions obtained in the first iteration of the taxonomy development process were derived. Similar characteristics were combined into a single characteristic. For example, the chemical, food, automotive, steel, and other industries were combined into the *manufacturing industry* characteristic. The production of various hardware components used in the fields of sensor technology, electronics, networking, and machines was combined to *hardware development*. Data analysis and the digital representation of these data were combined into *condition monitoring*. During this empirical iteration, it was found that the taxonomy required a dimension called *clients*, which represents important characteristics missing in the *customer segment* dimension. The IoT and PdM business models can be differentiated by considering the various segments that the customers of the companies operate in. For example, some companies have customers who sell to other companies (*B2B2B*). Further, it was recognized that another differentiation was the technical layer in which the offerings of the companies reside. This technical layer consists of four levels: recording, processing, handling, and analysis of data. This technical layer refers to the layers of IoT architecture models described by, e.g., Chen (2013) or Turber et al. (2014). Such models are used to describe the different prerequisites that are necessary for machine-to-machine communication. In the business model context, the technical layer is used to categorize layers in which the solutions offered by the companies reside. The end conditions of the taxonomy were not reached due to the newly identified dimensions and characteristics. Furthermore, there was a significant change in the taxonomy.

For the third iteration, the empirical-to-conceptual approach was used again. A random sample of 20 other companies was examined to check whether the dimensions and characteristics identified in the first two iterations were stable enough (i.e., sufficient number and chosen meaningfully). This iteration combined the provision of infrastructures, platforms and software in a public cloud. The development of algorithms for the analysis of data sets and their representation and the development of programs used for data security, encryption, and secure communication via the Internet are based on the development of mathematical algorithms. These are written programs and thus are similar to each other. Therefore, the newly identified characteristic *development of security software* was added to the *software development* characteristic. Customer segments such as the military and healthcare organizations were combined into *high-security areas*. The largest changes during this iteration occurred in

the revenue stream dimension. It was found that the revenue stream dimension was not entirely accurate for this taxonomy, as the definition provided by Osterwalder and Pigneur (2010) was too imprecise for our taxonomy. To describe this dimension more precisely, it was renamed the *payment model*. In addition, payment models consisting of a combination of several models, such as *one-time payment*, *project*-based payment, and/or subscription (payment on a *time basis*), were combined in the *hybrid* characteristic. A new type of payment *usage basis*, which is similar to *time basis* was identified and added to the taxonomy. In contrast to *time basis*, *usage basis* refers to billing based on the use of a particular resource (e.g., used computing capacity). The sales channel dimension was renamed into *deployment channel*. It was found that a better differentiator is how a customer accesses a service than how it is purchased. In the third iteration, there was also a significant change in the taxonomy, indicating that the end conditions had not been met.

An additional 30 companies were examined using the empirical-to-conceptual approach. Large companies such as Bosch Rexroth or National Instruments could not be assigned to a single key activity because they are active in many different areas (consulting, hardware development, software development, etc.). Accordingly, the activities of such companies were combined into the *universal range* characteristic. Furthermore, the newly identified customer segments logistics, aviation and railway were combined into the *logistics/transport industry* segment, as these segments are similar in their scope. An additional customer segment, a combination of *manufacturing industry + energy sector,* was identified and added to the taxonomy. In the fourth iteration, there was no

significant change in the taxonomy, but some characteristics were added. Thus, not all end conditions of taxonomy development were fulfilled.

Finally, the 53 remaining companies in the sample were examined. During this investigation, no further dimensions or characteristics were added or changed. Thus, according to Nickerson et al. (2013), the five subjective and eight objective end conditions of the taxonomy development were considered to be fulfilled. Formally, the final taxonomy was exactly the same as the taxonomy after the fourth iteration step.

The taxonomy developed for predictive maintenance business models

In the following, we present the final version of the developed taxonomy. Table 1 shows the dimensions in the first column and the characteristics in the respective rows. The first dimension *key activity* describes what the company primarily does according to its business model (Osterwalder et al. 2005; Osterwalder and Pigneur 2010). The second dimension, *value promise,* describes how customer needs are satisfied and customer problems are solved (Osterwalder et al. 2005, Osterwalder and Pigneur 2010). The *payment model* dimension describes how the performance of a PdM provider is measured and billed. For example, the *project* characteristic indicates that the company is paid after the execution of a defined project. This characteristic is therefore likely to be found frequently in consulting firms. On the other hand, the *time basis* characteristic indicates that the company bills for a certain period, for example, for the use of a cloud platform for

Table 1 Developed taxonomy

Dimensions	Characteristics		
Key activities	1) Hardware development	2) Software development	3) Consulting
	4) Edge computer development	5) Provision of a public cloud	6) Hardware retailing
	7) Universal range	8) Provision of an application platform	
Value promise	1) All-in-one solution	2) Condition monitoring	3) Connectivity
	4) Automation	5) Forecasting	6) Data security
	7) Data storage + software development tools		
Payment model	1) One-time sales	2) Time basis	3) Project
	4) Usage basis	5) Hybrid	
Deployment channel	1) Physical	2) www	3) Physical + www (cloud)
	4) www (cloud) + API	5) www (cloud)	6) Physical + www (cloud) + API
Customer segment	1) Manufacturing industry	2) Energy sector	3) No industry focus
	4) High-security areas	5) Manufacturing industry + energy sector	6) Manufacturing industry + logistics/transport industry
Clients	1) B2B	2) B2B + B2B2B	3) B2B + state
Information layer	1) Application and services	2) Information handling	3) Information delivering layer
	4) Object sensing and information gathering layer	5) Multiple	

one month. However, it is also possible to pay according to actual usage (*usage basis*), for example, according to the computing power used. The *deployment channel* dimension describes how a product or service is provided to the customer. To distinguish the companies according to their *customer segments*, this dimension describes the segment in which most of its customers operate (Osterwalder et al. 2005, Osterwalder and Pigneur 2010). The *clients* dimension describes the type of customer that purchases the service. The last dimension *information layer* represents the level at which the service is provided. This dimension is based on the IoT architecture model proposed by Chen (2013). The architecture model characterizes the different components required for a PdM application. The definitions of the characteristics of each dimension are provided in Table 6 of the Appendix.

Application of the taxonomy

Mapping of the sample

To show the applicability of the taxonomy, we assigned all 113 companies in the dataset to their respective characteristics. The websites for all these companies were used as the basis for the mapping. For companies that were identified on the fair website, information from the interviews with company representatives was used as supplementary material. If the Crunchbase database served as a source, information could be obtained from the short description provided. The process of mapping the characteristics of the companies in the data set was divided among four of the authors. For borderline companies, additional authors reviewed the information. To ensure that the mapping process used by the authors was as similar as possible, 10% of the companies were processed again by the authors. The level of agreement was measured using Fleiss' (1971) kappa coefficient. For this, we calculated the average agreement of the researchers for all dimensions for every company. Fleiss' (1971) kappa coefficient was 0.64, which, according to Landis and Koch (1977), corresponds to "substantial agreement". Therefore, it can be assumed that there was no significant bias in the results caused by using multiple authors for the mapping process. Table 2 shows the distribution of each dimension.

For the *key activity* dimension, the different characteristics are relatively evenly distributed with the exception of *hardware retailing* and *public cloud offering*. This observation might be related to the fact that providers of such services do not explicitly advertise PdM and are therefore not represented in our data set. The *value promise* dimension is dominated by *condition monitoring*, *forecasting* and *all-in-one solutions*. *Data security* is only weakly represented, which could be because companies that specialize in security do not explicitly advertise PdM services. For the *payment model*

dimension, payment on a usage basis still plays a subordinate role. Most companies use *hybrid* forms of payment. In the *deployment channel* dimension, the *physical* provision of products plays a major role as do *cloud and software* offerings. Most companies do not focus on an explicit sector. If companies have a focus, it is primarily on the manufacturing *industry*. PdM providers primarily prefer to do business in a *B2B* environment. An explicit focus on *state* or government organizations is rare. For the *information layer* dimension, about one-third of the companies provided services in several layers. Many companies provide services in the *application and services* layer. The fewest services are provided in the *information delivering* layer.
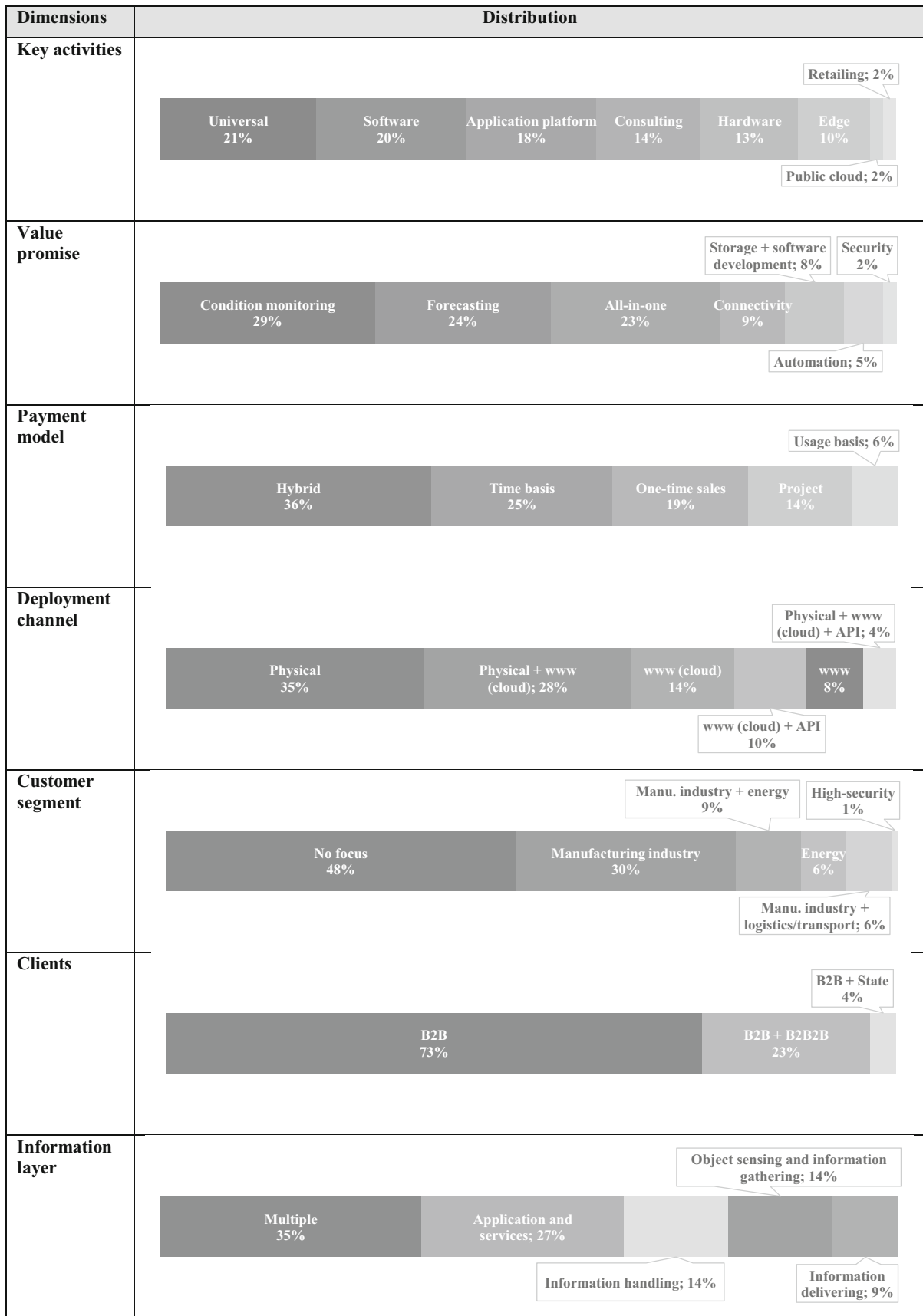
Business model clusters

To gain a better understanding of the PdM market, we created archetypes of the PdM environment equivalent to taxonomies developed for other business models (Gimpel et al. 2017; Eickhoff et al. 2017). To this end, we conducted a cluster analysis. Since our requirements are almost identical to those proposed by Gimpel et al. (2017), we also used Ward's (1963) algorithm for clustering. We also needed an algorithm that clusters our data, but the number of clusters was unknown. Ward's (1963) algorithm is a hierarchical cluster algorithm that not forms a predefined number of clusters, but all possible clusters are formed (Gimpel et al. 2017; Backhaus et al. 2011). This is done by calculating the differences between all objects. The differences are expressed as distances. We used the Sokal and Michener (1958) matching coefficient as the distance measure. Gimpel et al. (2017) indicate that there are various algorithms that can be used to determine a suitable number of clusters. We also used these algorithms, but depending on the algorithm, the results can be quite different. The results of the algorithms are provided in Table 8 in the Appendix. Because the algorithms produced very different results, we first created a graphical representation of the results of the Ward algorithm (Täuscher and Laudien 2018).

Figure 2 shows that three different clusters can be identified at the upper two branches. After looking at the companies assigned to the respective clusters, the groups were assigned the following three labels: "Universal vendors", "Software and platforms", and "Hardware and consulting". This label indicates that the groups were not yet granular enough. The next branching led to four groups. Since the height and thus the distance of the groups are similar to those in the next two branches, we analyzed six groups (rather than four) in the next step. These groups are identified in Fig. 2 with six rectangles, which are highlighted by bold lines.

After analyzing the companies in the six different groups, we developed the following labels for the clusters: "Consulting", "Hardware development", "Platform provider", "Information manager", "Analytics provider", and "All-in-

Table 2 Distribution of the characteristics

Dimensions	Distribution
Key activities	Retailing; 2% Universal 21% / Software 20% / Application platform 18% / Consulting 14% / Hardware 13% / Edge 10% Public cloud; 2%
Value promise	Storage + software development; 8% — Security 2% Condition monitoring 29% / Forecasting 24% / All-in-one 23% / Connectivity 9% Automation; 5%
Payment model	Usage basis; 6% Hybrid 36% / Time basis 25% / One-time sales 19% / Project 14%
Deployment channel	Physical + www (cloud) + API; 4% Physical 35% / Physical + www (cloud); 28% / www (cloud) 14% / www 8% www (cloud) + API 10%
Customer segment	Manu. industry + energy 9% — High-security 1% No focus 48% / Manufacturing industry 30% / Energy 6% Manu. industry + logistics/transport; 6%
Clients	B2B + State 4% B2B 73% / B2B + B2B2B 23%
Information layer	Object sensing and information gathering; 14% Multiple 35% / Application and services; 27% Information handling; 14% — Information delivering; 9%

one". We then increased this to seven groups. However, this led to a deterioration of the cluster results. Therefore, we came to the conclusion that the use of six groups was the most reasonable. In addition to using hierarchical cluster algorithms, partitioning algorithms could be used for the final allocation of companies to the six clusters. Following the approach proposed by Hartmann et al. (2016), we considered using the k-means and k-medoids algorithms for creating the clusters. K-medoids do not react as strongly to outliers as the k-means algorithm, which is an advantage of using this process (Hartmann et al. 2016). The decision regarding the quality of the assignment of the two algorithms was made by analyzing the distribution of the characteristics in each group. As shown by Hartmann et al. (2016), the k-medoid algorithm led to better results, which are shown in Table 3.

Table 3 shows the clear differentiation of the formed groups in terms of the first dimension (*key activities*). At least 61% of each group was assigned to the same characteristic. We named the groups based on the *key activities* dimension. For archetypes four and six, we also included other dimensions in the names, especially the second and seventh dimension.

The *hardware development* group is mainly made up of firms that develop and sell hardware (D1) but includes some of the companies that develop *edge devices*. The main value proposition (D2) is *condition monitoring*, but *automation and connectivity* also play a role. For this group, business is mainly conducted through *one-time sales* (D3); therefore, the *deployment channel* (D4) is *physical*. Most companies do not have a

specific industry focus, but if there is one, it is the *manufacturing industry*. Customers mainly operate in the *B2B* segment. All companies that explicitly mention the *state* as a customer are part of the *hardware development* group. The majority of companies work in the *object sensing and information gathering* layer.

The *platform provider* group comprises vendors of application platforms (D1) with a focus on *forecasting* models (D2). This work is done using a *hybrid payment model* (D3). Since some vendors do this in combination with *consulting* services and special *hardware devices*, the *deployment channel* (D4) is both *physical* and via a *cloud* platform. In this group, there is a focus on the manufacturing *industry* (D5). The customers operate in the *B2B* environment (D6), and the companies mainly operate in the *application and services* layer (D7).

Various services are offered by the companies assigned to the *all-in-one* group (D1). Mainly, they provide an *all-in-one* solution (D2). These companies use a *hybrid* (D3) *payment model*, and the *deployment channel* is both *physical* and via *cloud* solutions (D4). There is no specific customer segment (D5), and the customers operate in the B2B environment (D6). These companies are active in all information layers (D7). The *all-in-one* group is the largest group in the data set.

In contrast, the *information manager* group is the smallest in the data set. This group mainly consists of companies that *develop edge devices*, but *software development* and *consulting* also play a role (D1). The most common *value promise* is *condition monitoring* (D2). The *payment model*
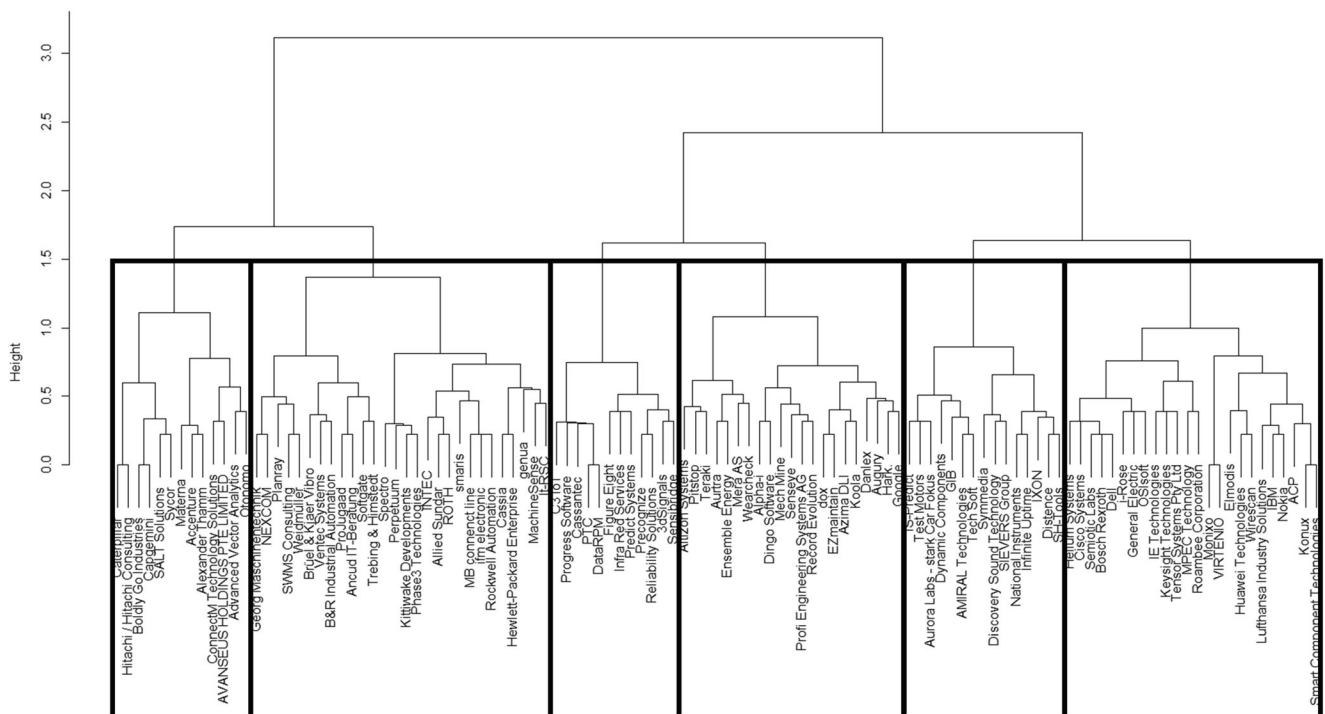

Fig. 2 Result of the Ward clustering visualized by a dendrogram

Table 3. Results of the cluster analysis. Note: Due to rounding inaccuracies, the sum of each column for each of the seven dimensions is not always exactly 100%

		Hardware development	Platform provider	All-in-one	Information manager	Consulting	Analytics provider
		1	2	3	4	5	6
D1	Provision of an application platform		64%	6%		7%	35%
	Edge computer development	17%		10%	67%		
	Hardware development	63%					
	Hardware retailing	8%					
	IT consulting			6%	17%	87%	
	Provision of a public cloud			3%			4%
	Software development	8%	21%	6%	17%	7%	61%
	Universal range	4%	14%	68%			
D2	Automation	17%	7%			7%	
	Data storage + software development tools	4%		13%	17%		13%
	All-in-one solution	8%		55%	17%	27%	9%
	Forecasting	8%	57%	3%	17%	20%	52%
	Data security	8%					
	Connectivity	13%	14%	13%		7%	
	Condition monitoring	42%	21%	16%	50%	40%	26%
D3	Usage basis		21%	6%		7%	4%
	Time basis		7%	19%	17%		87%
	One-time sales	83%		3%			
	Hybrid	13%	64%	71%	83%		9%
	Project	4%	7%			93%	
D4	Physical	83%		13%	33%	87%	4%
	Physical + www (cloud)	4%	36%	65%	50%		13%
	Physical + www (cloud) + API	4%		10%		7%	
	www	8%	14%	3%	17%		13%
	www (cloud)		36%	3%			43%
	www (cloud) + API		14%	6%		7%	26%
D5	Energy sector	8%	7%	3%	17%		9%
	High-security areas	4%					
	Manufacturing industry	29%	57%	16%	67%	33%	22%
	Manu. industry + energy sector	8%	14%	13%	17%	7%	
	Manu. industry + logistics/transport industry	4%	7%	6%		7%	9%
	No industry focus	46%	14%	61%		53%	61%
D6	B2B	71%	71%	74%	67%	67%	83%
	B2B + B2B2B	17%	29%	26%	33%	27%	17%
	B2B + state	13%				7%	
D7	Application and services		71%	6%		60%	43%
	Information delivering	17%		3%	33%		13%
	Information handling	4%	21%	13%		7%	30%
	Multiple	25%		74%	50%	33%	13%
	Object sensing and information gathering	54%	7%	3%	17%		

(D3) and *deployment channel* (D4) are the same as those of the *all-in-one* group (*hybrid and physically + www (cloud)*). In the *information manager* group, there is a focus on *industrial* companies (D5). This group's customers are not all *B2B* customers; one-third involve *B2B2B* relationships (D6). One-half of the companies are active in *multiple* information layers. One-third of the groups is engaged in *information delivering* (D7).

We call the fifth cluster the *consulting* group. Obviously, the *key activity* here is *consulting* (D1). The *value proposition* is *condition monitoring* (D2), and the payments are *project*-based (D3). Above all, this group provides consulting services. Since a consulting service cannot currently be provided automatically via a cloud or software, the *deployment channel* is primarily *physical* (D4). There is no special customer focus (D5), and the segment is *B2B* customers (D6). Consulting work is mainly done in the area of *application and services,* but one-third of the companies operate in *multiple* layers (D7).

The last group *analytics provider* mainly deals with *software development,* but the *provision of application platforms* also plays a role (D1). The *value promise* mainly focuses on the provision of *forecasts* (D2). The billing takes place on a *time basis* (D3) and the services are provided via a *cloud* (D4). There is no industry focus (D5), and services are provided to *B2B* customers (D6). The companies are mostly active in the *application and services* layer, but services are also provided in the *information handling* layer. The groups described are summarized in Table 4.

To validate the previous results and better understand the different archetypes, we use a visualization technique that allows us to locate each firm in a two-dimensional coordinate system. The two-dimensional representation of the multiple dimensions and characteristics can be easily interpreted and provides visual insights about the connections and relations among the groups. To develop the illustration, principal component analysis (PCA) could have been used to reduce the available information into two dimensions (Wang et al. 2016). When highly nonlinear dependencies appear in the data, other dimensionality reduction techniques such as autoencoders are superior to classical PCA (Wang et al. 2016). Since autoencoders are based on artificial neural networks that are trained to replicate the inputs and outputs, the network architecture can be designed to accurately learn the nonlinear relationships in the presented data. For a detailed description of this method, please see Hinton and Salakhutdinov (2006). Our autoencoder uses all available characteristics of the firms as input variables, and we encode categorical variables as multiple binary variables. The inputs are then passed forward through the network architecture of three hidden layers with 10, 2 and 10 neurons. Each layer is fully connected to the next layer. The third hidden layer is fully connected to the output of the

Table 4 Archetypes of PdM business models

	Archetype					
	1	2	3	4	5	6
Label	*Hardware development*	*Platform provider*	*All-in-one*	*Information manager*	*Consulting*	*Analytics provider*
Key activities	Hardware development	Provision of an application platform	Universal offer	Edge computer development	Consulting	Software development
Value promise	Condition monitoring	Forecasting	All-in-one solution	Condition monitoring	Condition monitoring	Forecasting
Payment model	One-time sales	Hybrid	Hybrid	Hybrid	Project	Time basis
Deployment channel	Physical	Physical + www (cloud)	Physical + www (cloud)	Physical + www (cloud)	Physical	www (cloud)
Customer segment	No industry focus	Manufacturing industry	No industry focus	Manufacturing Industry	No industry focus	No industry focus
Clients	B2B	B2B	B2B	B2B + B2B2B	B2B	B2B
Information layer	Object sensing and information gathering	Application and services	Multiple	Multiple & information delivering	Application and services	Application, services and information handling
Share in sample (113)*	21%	12%	27%	5%	13%	20%
Example company	Rockwell Automation	Test Motors	National Instruments	IXON	Hitachi Consulting	Senseye

*Due to rounding inaccuracy, the sum is not exactly 100%

autoencoder, which represents the input. Using backpropagation, the autoencoder learns the general rules that appear in the data through an iterative training process, which cause the output values to represent an accurate reconstruction of the input values. The encoding process is applied to every firm under study. Due to the architecture of two neurons in the second hidden layer, the autoencoder is forced to represent the input data by just two principal components that can be used to plot the firms in a two-dimensional space. Fig. 3 shows the resulting representation. Each firm is visualized by a dot in the coordinate system, while the different symbols indicate the affiliation of the firms based on k-medoid clustering.

Initially, three large groups can be identified in the visualization: In the range x < −0.2, in the quadrant x > 0, y < 0 and the group x, y > 0. In our analysis, we find that the composition of these three groups is approximately 80% consistent with the allocation of the three largest groups identified by the Ward algorithm, as shown in Fig. 2. The companies in the *all-in-one* group are represented in all areas, indicating that the generalists are not separated as an independent group in this representation but their business models have a different focus, which leads to the widespread allocation of such firms to the other clusters. Nevertheless, these companies usually distinguish themselves from the other groups by the first two dimensions, *key activities* and *value promise,* which justifies their assignment to a separate group. The companies in the *analytics and platform provider* groups are similar. This coincides with our experience in assigning the characteristics

and is shown in Table 4. Both groups are primarily concerned with the creation of *forecasts*, both operate in the *application and software* layer, and the companies in both groups sell either *software development*, the use of software or a *cloud platform*. The *consulting* and *hardware development* groups also seem to be similar. On closer inspection, however, it can be seen that the distances are much greater than between the *analytics* and *platform provider* groups. In both groups, *condition monitoring* is the *value promise* of the majority of companies and the *deployment channel* is *physical*. However, the *information layer* dimension indicates that the companies in the respective groups provide different services.

Discussion, implications, and recommendations

The developed taxonomy, which consists of seven dimensions, has enabled us to conduct cluster analysis. Cluster analysis is used to identify the archetypes of PdM business models that are currently utilized. The investigation of data-driven business models highlighted that the archetypes identified have similarities (Hartmann et al. 2016). The data source and the general aspects of the business model such as the key activities or value promises play a decisive role in data-driven business models (Engelbrecht et al. 2016). In our taxonomy, the data source plays a small role because for PdM, data are always obtained from the respective customers, and
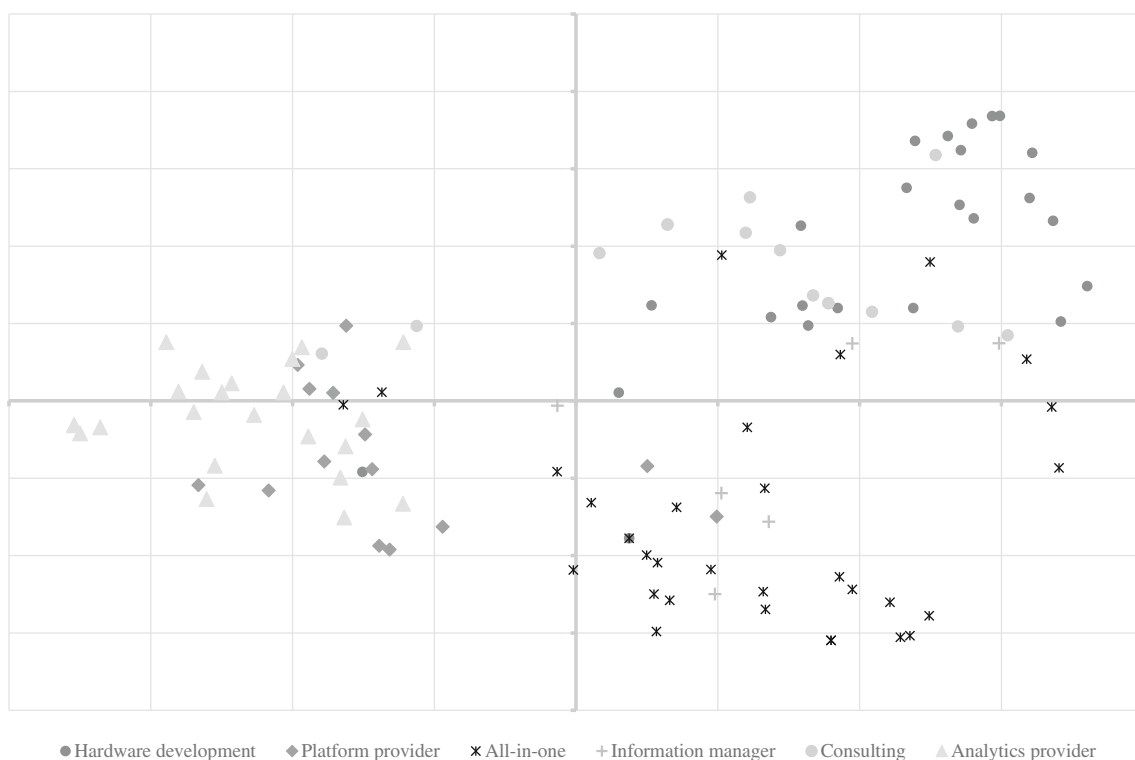

Fig. 3 Visualization of the clustering using an autoencoder method

therefore, there is little differentiation in the data sources for most PdM services. A new element of digital business models is needed to represent differences in a meaningful way. In the taxonomy developed here, the new dimension is the information layer. To develop this dimension, we used an IoT architecture model (Chen 2013). The architecture model shows the different components that are necessary for IoT or IIoT business models. This new dimensions may be useful for other, new business models, as they allow for a comparison of applications that are sometimes very complex.

We have identified six different archetypes. Among them are types that have already been described in this context and others that have been discussed rather rarely. In previous research on data-driven business models, there is a distinction between "data-aggregation-as-a-service" and "analytics-as-a-service" that is similar to the distinction between the archetypes *platform and analytics provider* identified in our study (Täuscher and Laudien 2018). *Analytics providers* analyze customer data and develop software for this purpose, while *platform providers* only provide the prerequisites for further analyses. An *analytics provider* works on the *application and services* layer and takes care of *information handling*. In addition to these archetypes, there are also hardware developers, consultants and generalists in the PdM environment. When considering the sample companies assigned to the *hardware development* archetype, we found companies (e.g., ROTH) that address problems such as repairing older machines that are not yet Internet-capable (retrofit). In these cases, hardware is needed to enable further analysis of PdM. This archetype has played a minor role in previous research and probably has a unique perspective of IIoT applications. In parallel, the presence of the *all-in-one* archetype supports the insight noted by Dijkman et al. (2015) that it is important for IoT business models to be convenient, usable and capable of "getting the job done" (Dijkman et al. 2015, p. 676). *All-in-one* offers seem to meet this demand. For example, National Instruments supplies software for analyzing collected data and various monitoring devices for recording different sensors. The existence of consulting firms shows that certain typed of PdM are complex and/or additional resources are needed.

Three major groups of business models are identified. These three groups were previously identified in the analysis using the Ward algorithm. The first group consists roughly of the analytics and platform providers, the second includes hardware development and in parts consulting, while the last group is mainly characterized by all-in-one providers. Following Bock and Wiener (2017), we can confirm that there are digital business models involving both "born-online" and "born-offline" companies. We find these types of firms in the *hardware development* ("born-offline"), *analytics* and *platform providers* ("born-online") archetypes. Specifically for PdM, there are all-in-one providers that combine both groups. These additional archetypes represent a difference to previous research. Thus far, the business models used by consultancies and all-in-one providers have not been mentioned in the context of data-driven business models. This finding also shows the need to look at real-world IoT or IIoT 4.0 applications. We suspect that previous research, which has primarily looked at B2C applications, has not been able to find these types of business models, as B2C applications are often less complex than B2B applications (Müller and Buliga 2019).

The taxonomy developed here can be used for the initial classification of a specific type of IoT business model, namely, a PdM business model. Many elements of the taxonomy are similar to those in other taxonomies. For example, other taxonomies often include the *value proposition* and *key activities* dimensions (e.g., Hartmann et al. 2016; Eickhoff et al. 2017; Täuscher and Laudien 2018). However, the respective characteristics and the *information layer* dimension are different. One important step was differentiating the sample companies according to the IoT layers to identify the differences between the individual PdM business models. Previous findings on Industry 4.0 business models are further specified by our taxonomy (Weking et al. 2018). We identify the characteristics of a specific type of Industry 4.0 business model. Since business models for PdM do not consist of only analytics services, it is necessary to consider analysis-based services (Rizk et al. 2018; Hunke et al. 2019), changes in hardware (Herterich et al. 2016), and platforms (Täuscher and Laudien 2018; Hodapp et al. 2019).

The analysis of a specific business model can identify the different components relevant for using the IoT or IIoT. These components might also be important for other IoT business models. For example, the structure of "pay as you drive" or "pay as you live" applications could be similar. For both of these applications, tariff reductions are granted based on careful driving or healthy living. Sensor data can be used to determine whether an insured drives carefully or lives a healthy life. It is also conceivable that there are similar business models, such as those used by sensor/device manufacturers and analysis providers. For these IoT applications, all layers of the IoT architecture must be included. However, in these IoT applications, a single vendor may offer services that reside in multiple layers, as the IoT application scenarios mentioned in prior research are often less complex. Thus, the distribution of the providers is different, although the basic structures are similar, meaning that all layers of an IoT architecture must be covered by one or more providers. Differentiating IoT or IIoT business models with the help of an architecture model facilitates the differentiation of the providers (Turber et al. 2014). Our taxonomy can serve as the basis for research on other IoT and IIoT applications.

The results of our research also have practical implications. The taxonomy developed in this study provides information about the dimensions that are important when

considering PdM business models. This taxonomy can be used by companies to classify their own business model and the models of other companies thereby facilitating a comparison. The archetypes identified in this study can facilitate these comparisons. Companies can use these archetypes to identify whether their PdM services are rare or common on the market. This assessment helps companies engage in specialization. However, companies that simply want to use PdM services can benefit from the results of our research. According to Dijkman et al. (2015), it is important for companies employing IoT business models to understand how others make money in the ecosystem. The results of our study contribute to the understanding of PdM business models, and thus, companies can optimize their networks or ecosystems. For example, depending on their individual application, companies can decide the extent to which they want to incorporate various elements and, if so, which parts of these elements. Firms can even outsource the complete implementation to one or more suppliers. Our taxonomy is helpful because it identifies the options that are available and the elements that are necessary to consider for implementation decisions.

Our article identifies why it is necessary to conduct research in an Industry 4.0 environment. For example, the retrofit of hardware plays an important role, increased networking is necessary for value creation, standardization plays an important role, and other payment and communication options are available in the industry (Müller and Buliga 2019).

The *hardware development*, *analytics provider,* and *all-in-one* archetypes are the most strongly represented in the data sample of this study. In contrast, there are few companies represented by the *information manager* archetype. There are three possible reasons for this: First, currently, there are few companies that pursue such a business model, which means that there is still a gap in the market that also represents growth opportunities for companies. Second, there are many companies that employ a similar business model, but they do not promote services such as "condition monitoring" and "predictive maintenance". For such companies, the use of more extensive or modified marketing strategies may be useful. Third, there is no demand for these types of services because companies do not need specific solutions or use their own solutions.

The use of an autoencoder and the subsequent dendrogram provided another way to visualize the similarities of the sample companies. In our case, the dendrogram was better suited for identifying a meaningful number of groups, while the two-dimensional chart better represents the distances and overlaps of the respective groups. Based on our experience, using a combination of both methods is an efficient approach for the formation and interpretation of archetypes.

Limitations and further research

In addition to the knowledge gained from this research, the limitations of this research must also be mentioned. The taxonomy depends on the definition used for PdM. Based on the existing literature, we comprehensively defined PdM and regarded, for example, "condition monitoring" as a part of PdM. If PdM is defined differently, a different taxonomy would result. The size of the sample used in this study is limited, and this is particularly evident in results for the cluster analysis, specifically for the *information manager*. Only 5% of the sample was assigned to this cluster, which makes the cluster very small and therefore the results associated with this cluster are difficult to interpret. Further research with more data should be conducted in this area; the use of more data will better articulate the description and thus improve the analysis of this archetype.

The sample includes only companies that can be found when searching for the terms "predictive maintenance" or "condition monitoring". There may be companies that provide PdM-like services but do not explicitly refer to them as PdM and thus these companies would not be found in our search. We tried to address this issue by using overview online articles to identify companies for this investigation. However, many companies in these articles explicitly use the term PdM. Additionally, some companies only roughly explain what services they provide. We tried to address this problem by having several authors analyze the borderline companies.

An autoencoder procedure was employed to develop a visualization of the business models identified by the taxonomy developed in this study. This analysis has led to valuable insights and made the PdM market transparent. Further research is needed to gain more experience with this process and to determine whether it can provide useful insights for other studies.

Our research provides a snapshot in time. The identified archetypes must be checked again after a certain period of time because the market is dynamic and changes may occur. Furthermore, new technologies have the potential to significantly change the market situation. Thus, the taxonomy will also change over time. Additional characteristics may be added, and it may be necessary to consider other dimensions. However, our taxonomy provides a starting point for further development.

Conclusions

In this study, we presented a taxonomy for the classification of PdM business models. Our taxonomy forms a basis for the classification of different providers of PdM solutions. To create the taxonomy, we examined a data set of 113 companies. Next, based on the taxonomy, we analyzed the business models. Using cluster analysis, we examined which archetypes of PdM business models currently exist and identified

six different archetypes. Our analysis of the archetypes showed that the general business model dimensions of the business model canvas (Osterwalder and Pigneur 2010) and the consideration of IoT architecture (Chen 2013) are important to differentiate the business models. PdM services are provided in all four layers of the architecture, whereby the information delivering layer is underrepresented in the sample companies. Although a wide variety of PdM services are provided in the market, we showed that PdM business models can be divided into six archetypes. Prior research on data-driven business models such as platforms and analytics providers have identified some archetypes found. In contrast, research has not yet described hardware development, all-in-one and consulting business models that operate in an IoT or IIoT environment. These insights offered by this study increase our understanding of PdM business models, both in theory and practice. Additionally, compared to general research on data-driven business models, hardware development is also important. In the B2B environment examined in this study, the retrofit of machines represents an important aspect of IIoT applications. Our research on the Industry 4.0 environment shows that the elements identified in research on data-driven, platform, analytics, and IoT business models can be found bundled together in practice when focusing on one specific IIoT application.

We used an autoencoder to visualize the identified PdM business models to better understand the different archetypes and their relationships. In combination with a dendrogram, this visualization enabled an analysis of the clusters that were identified. This procedure can be used in other studies and contribute to an efficient analysis.

Appendix

Content of the interviews

The following points were addressed in the interviews:

- Company name
- Position in the company (of the interviewee)
- Key activities
- Payment model
- Sales channel
- Customers
- Customer segment
- Value proposition for the customer
- Return of Invest for the customer
- Opinion regarding the expected market development

Iterations and end conditions

Table 5 Summary of fulfilled end conditions per iteration based on Nickerson et al. (2013)

Iteration					Ending conditions
1. con.*	2. emp.*	3. emp.*	4. emp.*	5. emp.*	
	x	x	x	x	Concise
	x	x	x	x	Robust
			x	x	Comprehensive
x	x	x	x	x	Extendible
				x	Explanatory
				x	All objects or a representative sample of objects have been examined
x	x	x	x	x	No object was merged with a similar object or split into multiple objects in the last iteration
	x	x	x	x	At least one object is classified under every characteristics of every dimension
				x	No new dimensions or characteristics were added in the last iteration
				x	No dimensions or characteristics were merged or split in the last iteration
x	x	x	x	x	Every dimension is unique and not repeated (i.e., there is no dimension duplication)
	x	x	x	x	Every characteristic is unique within its dimension (i.e., there is no characteristic duplication within a dimension)
	x	x	x	x	Each cell (combination of characteristics) is unique and is not repeated (i.e., there is no cell duplication)

* con. = conceptual, emp. = empirical

Definitions of the characteristics used

Table 6 Dimensions, their characteristics and their definitions

Dim.	Characteristic	Definition
D1 – Key activities	1) Hardware development	The development and manufacture of technical machine elements.
	2) Software development	The development/adaptation of programs for data processing systems.
	3) Consulting	Advising companies on the design, implementation and improvement of processes and solutions.
	4) Edge computer development	The development of systems for decentralized data acquisition/data processing at the edge of the network (also called "fog computing"). Can occur in combination with the use of a cloud. Can include both software and hardware development, but must involve a clear focus on edge computing.
	5) Provision of a public cloud	The provision of a computing and/or storage infrastructure accessible via the Internet.
	6) Hardware retailing	Purchases hardware components from various smaller manufacturers and distributes them to larger companies (occurs mainly in Asia; smaller, less well-known manufacturers use these retailers to sell their products globally).
	7) Universal range	Broadly based businesses with multiple key areas of activity.
	8) Provision of an application platform	A framework of services on which applications depend for standard operations. The platform includes operating systems, execution services, data services, cloud services and development tools.
D2 – Value promise	1) All-in-one solution	Complete software and hardware solutions from sensors to data storage to data analysis.
	2) Condition monitoring	The storage, analysis and display of machine data in real time (data must be provided).
	3) Connectivity	The provision of hardware and software components for setting up systems (e.g., routers, network cables, etc.).
	4) Automation	The provision of hardware components that enable the transfer of functions of the production process from humans to artificial systems (e.g., sensors). Components that enable a "retrofit".
	5) Forecasting	The forecasting of machine or component lifetimes or of the lifetimes of a machine part (e.g., section, component).
	6) Data security	The provision of security technology for both hardware and software (e.g., fire-resistant hardware components, encryption programs, etc.) for the implementation of predictive maintenance.
	7) Data storage + software development tools	The provision of large amounts of disk space and tools to create, debug, diagnose, and manage software.
D3 – Payment model	1) One-time sales	The product/service is paid for only once.
	2) Time basis	The product/service is paid for based on its usage period or at regular intervals (e.g., subscription or license for one year).
	3) Project	The product/service is paid for within the scope of a project, and after the project no further costs are charged for the service provided or for owning the developed output.
	4) Usage basis	The product/service is paid for on the basis of the amount of services used, the number of uses, the computing needs, etc.
	5) Hybrid	The combination of two or more payment models.
D4 – Deployment channelD4 – Dep	1) Physical	The provision of the product/service/hardware takes place physically (e.g., by implementing/installing software or hardware on site, consulting).
	2) www	The product/service can be downloaded or used via the Internet.
	3) Physical + www (cloud)	The combination of the physical provision of products/services (see description "physical") and use of services on a cloud accessible via the Internet.
	4) www (cloud) + API	Use of the product/service on a cloud accessible via the Internet. A programming interface is also provided.
	5) www (cloud)	Use of the product/service on a cloud accessible via the Internet.
	6) Physical + www (Cloud) + API	A combination of the abovementioned characteristics.
D5 – Customer segment		A more detailed definition of the segments is not necessary.
D6 – Clients	1) B2B	The products, services or products are sold to other companies. No sales to end customers.
	2) B2B + B2B2B	The combination of B2B (see above) and business-to-business-to-business (B2B2B). B2B2B is a B2B activity in which the customer of the focal company sells platform services to other companies (e.g., a company rents platform services and then sells them as services in addition to its product).
	3) B2B + State	Companies sell products/services to customers in the public sector (e.g., the military). This occurs in combination with B2B (see above).
D7 – Information layer	1) Application and services	Applications and services that use the acquired data (e.g., sensor data), e.g., for an analysis or forecast of the future deterioration of a machine (Chen 2013).
	2) Information handling	The processing of data and/or provision of computing capacity (Chen 2013).
	3) Information delivering layer	The transport and/or networking of information (Chen 2013).
	4) Object sensing and information gathering layer	The provision of sensors, data extraction and/or collection of information (Chen 2013).
	5) Multiple	The activities take place on more than one layer.

Predictive maintenance companies sample

Table 7 Company sample with name of the company, website, and source

Company	Website	Source
3dSignals	http://www.3dsig.com	Crunchbase Predictive Maintenance
Accenture	https://www.accenture.com/us-en/service-accenture-corrosion-management-services	Website List
ACP	https://www.acp.de	Hannover Industrie Messe 2018
Advanced Vector Analytics	http://www.ava-labs.com/	Crunchbase Condition Monitoring
Alexander Thamm	https://www.alexanderthamm.com/de/	Website List
Allied Sundar	http://www.sundar.com.tw/	Hannover Industrie Messe 2018
Alpha-i	http://alpha-i.co/	Crunchbase Predictive Maintenance
Altizon Systems	http://altizon.com/	Crunchbase Webpage Search
AMIRAL Technologies	https://www.amiraltechnologies.com/en/	Hannover Industrie Messe 2018
Ancud IT-Beratung	https://www.ancud.de/	Hannover Industrie Messe 2018
Augury	http://www.augury.com	Crunchbase Predictive Maintenance
Aurora Labs - stark Car Fokus	https://www.auroralabs.com/	Crunchbase Webpage Search
Aurtra	https://www.aurtra.com.au	Crunchbase Condition Monitoring
AVANSEUS HOLDINGS PTE LIMITED	http://www.avanseus.com/	Crunchbase Predictive Maintenance
Azima DLI	http://www.azimadli.com	Crunchbase Condition Monitoring
B&R Industrial Automation	https://www.br-automation.com/	Hannover Industrie Messe 2018
Boldly Go Industries	https://www.boldlygo.de	Hannover Industrie Messe 2018
Bosch Rexroth	https://www.boschrexroth.com/en/xc/service/industrial-applications/predictive-maintenance/predictive-maintenance-2	Website list
Brüel & Kjaer Vibro	https://www.bkvibro.com	Hannover Industrie Messe 2018
C3 IoT	https://c3iot.ai/	Crunchbase Webpage Search
Capgemini	https://www.capgemini.com	Hannover Industrie Messe 2018
Cassantec	https://casantec.com/	Website List
Cassia	https://www.cassianetworks.com/	Hannover Industrie Messe 2018
Caterpillar	https://www.cat.com/de_DE/support/maintenance/condition-monitoring.html	Website List
Cisco Systems	https://www.cisco.com/c/en/us/solutions/internet-of-things/overview.html	Website List
ConnectM Technology Solutions	http://www.connectm.com	Crunchbase Condition Monitoring
Danlex	http://www.danlex.com	Crunchbase Predictive Maintenance
DataRPM	http://www.datarpm.com/	Crunchbase Webpage Search
Dell	http://www.dell.com/en-us/work/learn/internet-of-things-solutions#Why-choose-Dell?	Website List
Dingo Software	http://www.dingo.com	Crunchbase Predictive Maintenance
Discovery Sound Technology	http://www.discoverysoundtechnology.com	Crunchbase Predictive Maintenance
Distence	https://www.distence.fi/de	Hannover Industrie Messe 2018
dox	http://dox.tech	Crunchbase Predictive Maintenance
Dynamic Components	http://www.dynamic-components.de	Crunchbase Predictive Maintenance
Elmodis	http://www.elmodis.com/	Crunchbase Webpage Search
Ensemble Energy	http://www.ensembleenergy.ai	Crunchbase Predictive Maintenance
EZmaintain	https://www.ezmaintain.com	Crunchbase Predictive Maintenance
Figure Eight	https://www.figure-eight.com/	Crunchbase Webpage Search
General Electric	https://www.ge.com/digital/sites/default/files/Predix-from-GE-Digital-Overview-Brochure.pdf	Website List
genua	https://www.genua.de	Hannover Industrie Messe 2018
Georg Maschinentechnik	http://www.georg-maschinentechnik.de/	Hannover Industrie Messe 2018
GIB	https://www.gibmbh.de/en/	Hannover Industrie Messe 2018
Google	https://cloud.google.com/	Hannover Industrie Messe 2018
Hark.	https://harksys.com	Crunchbase Predictive Maintenance
Helium Systems	https://www.helium.com/solutions/manufacturing	Website List
Hewlett-Packard Enterprise	https://www.hpe.com/de/de/solutions/industrial-internet-of-things.html	Hannover Industrie Messe 2018
Hitachi / Hitachi Consulting	https://www.hitachiconsulting.com/solutions/hitachi-predictive-maintenance.html	Website List
Huawei Technologies	https://e.huawei.com/en/solutions/business-needs/enterprise-network/agile-iot/elevators-connection	Website List
IBM	https://www.ibm.com/us-en/marketplace/ibm-predictive-maintenance-optimization	Website List
IE Technologies	http://www.ietechnologiesllc.com	Crunchbase Predictive Maintenance
ifm electronic	https://www.ifm.com/	Website List
Infinite Uptime	http://www.infinite-uptime.com/	Crunchbase Webpage Search
Infra Red Services	http://www.infraredservices.com.au	Crunchbase Condition Monitoring

Table 7 (continued)

Company	Website	Source
INTEC	https://www.intec-connectivity.com/	Hannover Industrie Messe 2018
i-Rose	http://www.i-rose.si	Crunchbase Predictive Maintenance
IS-Predict	http://www.ispredict.com/	Hannover Industrie Messe 2018
It-RSC	https://it-rsc.de/	Hannover Industrie Messe 2018
IXON	https://www.ixon.cloud/de	Hannover Industrie Messe 2018
Keysight Technologies	https://www.keysight.com/de/de/home.html	Website List
Kittiwake Developments	http://www.kittiwake.com	Crunchbase Condition Monitoring
Konux	https://www.konux.com/de/	Website List
Koola	http://www.koola.io	Crunchbase Predictive Maintenance
Lufthansa Industry Solutions	https://www.lufthansa-industry-solutions.com/de-en/	Hannover Industrie Messe 2018
MachineSense	https://machinesense.com/	Crunchbase Predictive Maintenance
Materna	https://www.materna.de/	Hannover Industrie Messe 2018
MB connenct line	https://www.mbconnectline.com/	Hannover Industrie Messe 2018
Mech Mine	https://www.mechmine.com/en/	Hannover Industrie Messe 2018
Mera AS	http://www.mera.no	Crunchbase Condition Monitoring
Monixo	http://www.monixo.com	Hannover Industrie Messe 2018
MPEC Technology	http://www.mpec.co.uk	Crunchbase Condition Monitoring
National Instruments	http://www.ni.com/de-de/innovations/industrial-machinery/condition-monitoring.html	Website List
NEXCOM	http://www.nexcom.com/	Hannover Industrie Messe 2018
Nokia	https://spacetimeinsight.com/asset-analytics/	Website List
OSIsoft	https://www.osisoft.com/iiot/	Website List
Otonomo	http://www.otonomo.io/	Crunchbase Webpage Search
Perpetuum	http://www.perpetuum.com	Crunchbase Condition Monitoring
Phase3 Technologies	http://www.phase3-tech.com/	Crunchbase Predictive Maintenance
Pitstop	https://www.pitstopconnect.com	Crunchbase Predictive Maintenance
Planray	http://www.planray.com/en/	Hannover Industrie Messe 2018
Precognize	http://www.precog.co	Crunchbase Predictive Maintenance
Predict Systems	http://predictsystems.com	Crunchbase Predictive Maintenance
Profi Engineering Systems AG	https://www.profi-ag.de/	Hannover Industrie Messe 2018
Progress Software	https://www.progress.com/solutions/cognitive-predictive-maintenance	Website List
ProJugaad	http://projugaad.com	Crunchbase Predictive Maintenance
PTC	https://www.ptc.com/en/products/iot/thingworx-platform/analyze	Website List
Record Evolution	https://record-evolution.de/	Hannover Industrie Messe 2018
Reliability Solutions	http://reliasol.pl/en/	Crunchbase Predictive Maintenance
Roambee Corporation	http://www.roambee.com	Crunchbase Condition Monitoring
Rockwell Automation	https://ab.rockwellautomation.com/Condition-Monitoring	Website List
ROTH	https://www.roth-gruppe.de/	Hannover Industrie Messe 2018
SALT Solutions	https://www.salt-solutions.de/	Hannover Industrie Messe 2018
Semiotic Labs	http://www.semioticlabs.com	Crunchbase Condition Monitoring
Senseye	http://www.senseye.io/	Hannover Industrie Messe 2018
Sensibridge	http://sensibridge.com/	Crunchbase Predictive Maintenance
SH-Tools	https://www.sh-tools.com/de/	Hannover Industrie Messe 2018
SIEVERS Group	https://www.sievers-group.com/	Hannover Industrie Messe 2018
smaris	https://www.smaris.cz/	Hannover Industrie Messe 2018
Smart Component Technologies	https://smartcomptech.com	Crunchbase Condition Monitoring
Softgate	https://www.soft-gate.de/	Hannover Industrie Messe 2018
Spectro	http://www.spectroinc.com	Crunchbase Condition Monitoring
SWMS Consulting	https://www.swms.de/consulting/	Hannover Industrie Messe 2018
Sycor	https://de.sycor-group.com/	Hannover Industrie Messe 2018
Symmedia	https://www.symmedia.de/	Hannover Industrie Messe 2018
Tech Soft	https://www.techsoft.at/	Hannover Industrie Messe 2018
Tensor Systems Pty Ltd	http://www.tensorsystems.com	Crunchbase Condition Monitoring
Teraki	http://www.teraki.com	Crunchbase Predictive Maintenance
Test Motors	http://www.testmotors.com	Crunchbase Predictive Maintenance
Trebing & Himstedt	https://www.t-h.de/	Hannover Industrie Messe 2018
Ventec Systems	http://www.ventech-systems.com	Crunchbase Condition Monitoring
VIRTENIO	https://www.virtenio.com/de/	Hannover Industrie Messe 2018
Wearcheck	http://www.wearcheck.co.za	Crunchbase Condition Monitoring
Weidmüller	https://www.weidmueller.de/	Hannover Industrie Messe 2018
Wirescan	http://www.wirescan.no	Crunchbase Condition Monitoring

Comparison of different algorithms for selecting the cluster amount

Table 8 Recommended amount of clusters of different algorithms

Measure by	Recommended amount of clusters
Ball and Hall (1965)	3
Caliński and Harabasz (1974)	2
Davies and Bouldin (1979)	12
Dunn (1974)	2
Frey and Van Groenewoud (1972)	1
Halkidi et al. (2000)	8
Hartigan (1975)	3
Hubert and Levin (1976)	9
Krzanowski and Lai (1988)	7
McClain and Rao (1975)	2
Milligan 1981)	12
Rousseeuw (1987)	9
Tibshirani et al. (2001)	2

References

Afuah, A., & Tucci, C. L. (2001). *Internet business models and strategies: Text and cases*. New York: MacGraw-Hill/Irwin.

Aivaliotis, P., Georgoulias, K., & Chryssolouris, G. (2017). A RUL calculation approach based on physical-based simulation models for predictive maintenance. *Proceedings of the International Conference on Engineering, Technology and Innovation*, Madeira, Portugal, June 25-28 (pp. 1243–1246). https://doi.org/10.1109/ice.2017.8280022.

Alt, R., & Zimmermann, H. D. (2001). Preface: Introduction to special section–business models. *Electronic Markets, 11*(1), 3–9. https://doi.org/10.1080/713765630

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2011). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Germany, Springer: Berlin/Heidelberg.

Baidya, R., & Ghosh, S. K. (2015). Model for a predictive maintenance system effectiveness using the analytical hierarchy process as

analytical tool. *IFAC-PapersOnLine, 48*(3), 1463–1468. https://doi.org/10.1016/j.ifacol.2015.06.293

Bailey, K. D. (1984). A three-level measurement model. *Quality and Quantity, 18*(3), 225–245. https://doi.org/10.1007/bf00156457

Barbera, F., Schneider, H., & Kelle, P. (1996). A condition based maintenance model with exponential failures and fixed inspection intervals. *Journal of the Operational research Society, 47*(8), 1037–1045. https://doi.org/10.1057/jors.1996.130

Bock, M., Wiener, M. (2017). Towards a taxonomy of digital business models – Conceptual dimensions and empirical illustrations. *Proceedings of the 38th International Conference on Information Systems*, Seoul, South Korea.

Borgi, T., Hidri, A., Neef, B., & Naceur, M. S. (2017). Data analytics for predictive maintenance of industrial robots. *Proceedings of the International Conference on Advanced Systems and Electric Technologies,* Hammamet, Tunisia, (pp. 412–417). https://doi.org/10.1109/aset.2017.7983729.

Brousseau, E., & Pénard, T. (2007). The economics of digital business models: A framework for analyzing the economics of platforms. *Review of Network Economics, 6*(2), 81–114. https://doi.org/10.2202/1446-9022.1112

BSI (British Standards Institution) (2010). Maintenance – Maintenance terminology. BS en 13306:2010. Milton Keynes, BSI.

Cachada, A., Barbosa, J., Leitño, P., Gcraldcs, C. A., Deusdado, L., Costa, J., Teixeira, C., Teixeira, J., Moreira, A. H. J., Moreira, P. M., & Romero, L. (2018). Maintenance 4.0: Intelligent and predictive maintenance system architecture. P*roceedings of the 23rd International Conference on Emerging Technologies and Factory Automation*, Torino, Italy, (pp. 139–146).

Chen, M. (2013). Towards smart city: M2M communications with software agent intelligence. *Multimedia Tools and Applications, 67*(1), 167–178. https://doi.org/10.1007/s11042-012-1013-4

Chu, C., Proth, J. M., & Wolff, P. (1998). Predictive maintenance: The one-unit replacement model. *International Journal of Production Economics, 54*(3), 285–295. https://doi.org/10.1016/s0925-5273(98)00004-8

Daily, J., & Peterson, J. (2017). Predictive maintenance: How big data analysis can improve maintenance. In K. Richter & J. Walther (Eds.), *Supply Chain Integration Challenges in Commercial Aerospace* (pp. 267–278). Cham: Springer. https://doi.org/10.1007/978-3-319-46155-7_18.

Darwanto, D., Hamdani, D., Hariyanto, D. D., & Karyawan, O. H. (2012). Partial discharge analysis for predictive maintenance of generator of geothermal power plant. *Proceedings of the International Conference on Condition Monitoring and Diagnosis*, Bali, Indonesia (pp. 740–743). https://doi.org/10.1109/cmd.2012.6416252.

Dijkman, R. M., Sprenkels, B., Peeters, T., & Janssen, A. (2015). Business models for the Internet of Things. *International Journal of Information Management, 35*(6), 672–678.

Eickhoff, M., Muntermann, J., & Weinrich, T. (2017). What do FinTechs actually do? A taxonomy of FinTech business models. *Proceedings of the International Conference on Information Systems*, Seoul, South Korea.

Engelbrecht, A., Gerlach, J., & Widjaja, T. (2016). Understanding the anatomy of data-driven business models – towards an empirical taxonomy. *Proceedings of the Twenty-Fourth European Conference on Information Systems*, Istanbul, Turkey.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378–382. https://doi.org/10.1037/h0031619

Gerloff, C., Cleophas, C. (2017). Excavating the treasure of IoT data: An architecture to empower rapid data analytics for predictive maintenance of connected vehicles. *Proceedings of the International Conference on Information Systems*, Seoul, South Korea.

Gimpel, H., Rau, D., & Röglinger, M. (2017). Understanding FinTech start-ups – a taxonomy of consumer-oriented service offerings. *Electronic Markets, 28*(3), 245–264. https://doi.org/10.1007/s12525-017-0275-0

Glass, R. L., & Vessey, I. (1995). Contemporary application-domain taxonomies. *IEEE Software, 12*(4), 63–76. https://doi.org/10.1109/52.391837

Gregor, S. (2006). The nature of theory in information systems. *MIS Quaterly, 30*(3), 611–642. https://doi.org/10.2307/25148742

Groba, C., Cech, S., Rosenthal, F., & Gossling, A. (2007). Architecture of a predictive maintenance framework. *Proceedings of the 6th International Conference on Computer Information Systems and Industrial Management Applications*, Elk, Poland (pp. 59–64).

Hartmann, P. M., Zaki, M., Feldmann, N., & Neely, A. (2016). Capturing value from big data – a taxonomy of data-driven business models used by start-up firms. *International Journal of Operations & Production Management, 36*(10), 1382–1406. https://doi.org/10.1108/ijopm-02-2014-0098

Herterich, M. M., Buehnen, T., Uebernickel, F., & Brenner, W. (2016). A taxonomy of industrial service systems enabled by digital product innovation. *Proceedings of the 49th Hawaii International Conference on System Sciences* (pp. 1236–1245). Hawaii: IEEE. https://doi.org/10.1109/hicss.2016.157.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information system research. *MIS Quaterly, 28*(1), 75–105. https://doi.org/10.2307/25148625

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Hodapp, D., Remane, G., Hanelt, A., & Kolbe, L. M. (2019). Business models for Internet of Things platforms: Empirical development of a taxonomy and archetypes. *Proceedings of the 14th International Conference on Wirtschaftsinformatik*, Siegen, Germany.

Holgado, M., & Macchi, M. (2014). Exploring the role of E-maintenance for value creation in service provision. *Proceedings of the International Conference on Engineering, Technology and Innovation* (pp. 1–10). Bergamo, Italy: IEEE. https://doi.org/10.1109/ice.2014.6871586.

Hui, D., Wang, T., Leng, Y., & Zhang, Y. (2008). Research of open condition monitor system based on the predictive maintenance. *Proceedings of the 4th International Conference on Natural Computation*, Jinan, China (Vol. 7, pp. 317–320). https://doi.org/10.1109/icnc.2008.730.

Hunke, F., Engel, C., Schüritz, R., & Ebel, P. (2019). Understanding the anatomy of analytics-based services–A taxonomy to conceptualize the use of data and analytics in services. *Proceedings of the 27th European Conference on Information Systems*, Stockholm & Uppsala, Sweden.

Khazraei, K., & Deuse, J. (2011). A strategic standpoint on maintenance taxonomy. *Journal of Facilities Management, 9*(2), 96–113. https://doi.org/10.1108/14725961111128452

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174. https://doi.org/10.2307/2529310

Last, M., Sinaiski, A., & Subramania, H. S. (2010). Predictive maintenance with multi-target classification models. *Proceedings of the Asian Conference on Intelligent Information and Database Systems*, Hue City, Vietnam (pp. 368–377). https://doi.org/10.1007/978-3-642-12101-2_38.

Mahadevan, B. (2000). Business models for internet-based E-commerce: An anatomy. *California Management Review, 42*(4), 55–69. https://doi.org/10.2307/41166053

Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. (2015). Unlocking the potential of the Internet of Things. *McKinsey Global Institute*, Retrieved from https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world. Accessed 15 Feb 2020.

Mattes, A., Schöpka, U., Schellenberger, M., Scheibelhofer, P., & Leditzky, G. (2012). Virtual equipment for benchmarking predictive maintenance algorithms. *Proceedings of the 2012 Winter Simulation Conference*, Berlin, Germany. https://doi.org/10.1109/wsc.2012.6465084.

Müller, J., Buliga, O. (2019). Archetypes for data-driven business models for manufacturing companies in industry 4.0. *Proceedings of the 40th International Conference on Information Systems*, Munich, Germany.

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems, 22*(3), 336–359. https://doi.org/10.1057/ejis.2012.26

Nickerson, R. C., Varshney, U., & Muntermann, J. (2017). Of taxonomies and taxonomic theories. *Proceedings of the twenty-third Americas Conference on Information Systems*, Boston, Massachusetts.

Osterwalder, A., Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, Gamechangers, and challengers.* John Wiley & Sons.

Osterwalder, A., Pigneur, Y., & Tucci, C. L. (2005). Clarifying business models: Origins, present, and future of the concept. *Communications of the Association for Information Systems, 16*(1). https://doi.org/10.17705/1cais.01601.

Paukstadt, U., Strobel, G., & Eicker, S. (2019). Understanding services in the era of the Internet of Things: A smart service taxonomy. *Proceedings of the 27th European Conference on Information Systems*, Stockholm & Uppsala, Sweden.

PricewaterhouseCoopers (2017). Predictive Maintenance 4.0. PricewaterhouseCoopers. Retrieved from https://www.pwc.nl/nl/assets/documents/pwc-predictive-maintenance-4-0.pdf [Accessed September 27, 2018].

Rizk, A., Bergvall-Kåreborn, B., & Elragal, A. (2018). Towards a taxonomy for data-driven digital services. *Proceedings of the 51st Hawaii International Conference on System Sciences*, Hawaii. https://doi.org/10.24251/hicss.2018.135.

Sipos, R., Fradkin, D., Moerchen, F., & Wang, Z. (2014). Log-based predictive maintenance. *Proceedings of the 20th International Conference on Knowledge Discovery and Data Mining*, New York (pp. 1867–1876).

Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Bulletin, 38*(2), 1409–1438.

Spendla, L., Kebisek, M., Tanuska, P., & Hrcka, L. (2017). Concept of predictive maintenance of production systems in accordance with industry 4.0. *Proceedings of the 15th International Symposium on Applied Machine Intelligence and Informatics*, Herl'any, Slovakia (pp. 405–410). https://doi.org/10.1109/sami.2017.7880343.

Susto, G. A., Beghi, A., & De Luca, C. (2012). A predictive maintenance system for epitaxy processes based on filtering and prediction techniques. *IEEE Transactions on Semiconductor Manufacturing, 25*(4), 638–649. https://doi.org/10.1109/tsm.2012.2209131

Susto, G. A., Schirru, A., Pampuri, S., Pagano, D., McLoone, S., & Beghi, A. (2013). A predictive maintenance system for integral type faults based on support vector machines: An application to ion implantation. *Proceedings of the IEEE International Conference on Automation Science and Engineering*, Madison, Wisconsin (pp.195–200). https://doi.org/10.1109/coase.2013.6653952.

Szopinski, D., Schoormann, T., & Kundisch, D. (2019). Because your taxonomy is worth IT: Towards a framework for taxonomy evaluation. *Proceedings of the European Conference on Information Systems*, Stockholm & Uppsala, Sweden.

Täuscher, K., & Laudien, S. M. (2018). Understanding platform business models: A mixed methods study of marketplaces. *European*

Management Journal, 36(3), 319–329. https://doi.org/10.1016/j.emj.2017.06.005

Turber, S., Vom Brocke, J., Gassmann, O., & Fleisch, E. (2014). Designing business models in the era of internet of things. *Proceedings of the International Conference on Design Science Research in Information Systems*, Miami, Florida (pp. 17–31). https://doi.org/10.22215/timreview807.

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems, 37*(9), 206–224. https://doi.org/10.17705/1cais.03709.

Wang, N., Sun, S., Si, S., & Li, J. (2009). Research of predictive maintenance for deteriorating system based on semi-Markov process. *Proceedings of the 16th International Conference on Industrial Engineering and Engineering Management*, Beijing, China (pp.899–903). https://doi.org/10.1109/icieem.2009.5344200.

Wang, Y., Yao, H., & Zhao, S. (2016). Auto-encoder based dimensionality reduction. *Neurocomputing, 184,* 232–242. https://doi.org/10.1016/j.neucom.2015.08.104

Wang, J., Zhang, L., Duan, L., & Gao, R. X. (2017). A new paradigm of cloud-based predictive maintenance for intelligent manufacturing. *Journal of Intelligent Manufacturing, 28*(5), 1125–1137. https://doi.org/10.1007/s10845-015-1066-0

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Webster, J. & Watson, T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), xiii-xxiii. http://www.jstor.org/stable/4132319.

Weking, J., Stöcker, M., Kowalkiewicz, M., Böhm, M., & Krcmar, H. (2018) Archetypes for industry 4.0 business model innovations. P*roceedings of the 24th Americas Conference on Information Systems (AMCIS)*, New Orleans, LA.

Whitmore, A., Agarwal, A., Da, Xu., & L. (2015). The Internet of Things - a survey of topics and trends. *Information Systems Frontiers, 17*(2), 261–274. https://doi.org/10.1007/s10796-014-9489-2

Zoll, M., Jäck, D., & Vogt, M. W. (2018). Evaluation of predictive-maintenance-as-a-service business models in the internet of things. P*roceeding of the IEEE International Conference on Engineering, Technology and Innovation*, Stuttgart, Germany,. https://doi.org/10.1109/ice.2018.8436272.

References of the algorithms

Ball, G. H., & Hall, D. J. (1965). *Isodata: A novel method of data analysis and pattern classification*. Menlo Park: Stanford Research Institute.

Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods, 3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 2,* 224–227. https://doi.org/10.1109/tpami.1979.4766909

Dunn, J. C. (1974). Well-separated clusters and optimal fuzzy partitions. *Journal of cybernetics, 4*(1), 95–104. https://doi.org/10.1080/01969727408546059

Frey, T., & Van Groenewoud, H. (1972). A cluster analysis of the D2 matrix of white spruce stands in Saskatchewan based on the maximum-minimum principle. *The Journal of Ecology, 873–886,*. https://doi.org/10.2307/2258571

Halkidi, M., Vazirgiannis, M., & Batistakis, Y. (2000). Quality scheme assessment in the clustering process. *European Conference on Principles of Data Mining and Knowledge Discovery* (pp. 265–276). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45372-5_26.

Hartigan, J. A. (1975). *Clustering algorithms*. New York: Wiley.

Hubert, L. J., & Levin, J. R. (1976). A general statistical framework for assessing categorical clustering in free recall. *Psychological Bulletin, 83*(6), 1072–1080. https://doi.org/10.1037/0033-2909.83.6.1072

Krzanowski, W. J., & Lai, Y. T. (1988). A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics, 44,* 23–34. https://doi.org/10.2307/2531893

McClain, J. O., & Rao, V. R. (1975). Clustisz: A program to test for the quality of clustering of a set of objects. *Journal of Marketing Research*, 456–460.https://www.jstor.org/stable/3151097.

Milligan, G. W. (1981). A Monte Carlo study of thirty internal criterion measures for cluster analysis. *Psychometrika, 46*(2), 187–199. https://doi.org/10.1007/bf02293899

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20,* 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63*(2), 411–423. https://doi.org/10.1111/1467-9868.00293