

## **UNDERSTANDING USAGE DATA-DRIVEN PRODUCT PLANNING: A SYSTEMATIC LITERATURE REVIEW**

**Meyer, Maurice (1);  
Wiederkehr, Ingrid (1);  
Koldewey, Christian (1);  
Dumitrescu, Roman (1,2)**

1: Heinz Nixdorf Institute, University of Paderborn;

2: Fraunhofer Institute for Mechatronic Systems Design IEM

### **ABSTRACT**

Cyber-physical systems (CPS) are able to collect huge amounts of data about themselves, their users, and their environment during their usage phase. By feeding these usage data back into product planning, manufacturers can optimize their engineering and decision-making processes. Despite promising potentials, most manufacturers still do not analyze usage data within product planning. Also, research on usage data-driven product planning is scarce. Therefore, this paper aims to identify the main concepts, advantages, success factors and challenges of usage data-driven product planning. To answer the corresponding research questions, a comprehensive systematic literature review is conducted. From its results, a detailed description of usage data-driven product planning consisting of six main concepts is derived. Furthermore, taxonomies for the advantages, success factors and challenges of usage data-driven product planning are presented. The six main concepts and the three taxonomies allow for a deeper understanding of the topic while highlighting necessary future actions and research needs.

**Keywords:** Data-Driven Product Planning, Big data, Machine learning, Early design phases, Product Lifecycle Management (PLM)

### **Contact:**

Meyer, Maurice  
Heinz Nixdorf Institute, University of Paderborn  
Advanced Systems Engineering  
Germany  
maurice.meyer@hni.uni-paderborn.de

**Cite this article:** Meyer, M., Wiederkehr, I., Koldewey, C., Dumitrescu, R. (2021) 'Understanding Usage Data-Driven Product Planning: A Systematic Literature Review', in *Proceedings of the International Conference on Engineering Design (ICED21)*, Gothenburg, Sweden, 16-20 August 2021. DOI:10.1017/pds.2021.590

## 1 INTRODUCTION

The megatrend digitalization has turned mostly mechanical and electrical products into complex systems integrating hardware, sensors, data storage, microprocessors, software and connectivity (Porter and Heppelmann, 2014). These so-called Cyber-Physical Systems (CPS) integrate computation and physical processes (Lee, 2008) enabling the collection, analysis and sharing of huge amounts of data about the product and its environment (Porter and Heppelmann, 2014). As a consequence, CPS turn their users into incessant generators of data (Erevelles *et al.*, 2016). By feeding these data from the usage phase back into product development and especially product planning, manufacturers can optimize their engineering and decision-making processes (Erwin *et al.*, 2015). As a result, these data enable the companies to turn their customers and users into a powerful source of innovation (Holler *et al.*, 2016a) and – consequently – to develop a competitive advantage (Porter and Heppelmann, 2015).

## 2 SCIENTIFIC BACKGROUND OF USAGE DATA-DRIVEN PRODUCT PLANNING

Product planning is the first activity in the product creation process. It aims at finding the success potentials of the future to create a promising product design in the form of a principle solution (Gausemeier *et al.*, 2011). The results of product planning are the products to be developed by the organization (Ulrich and Eppinger, 2016) and the corresponding requirements lists for the subsequent product development (Pahl *et al.*, 2007). The interconnections between product planning and product development are strong: For Gausemeier *et al.*, both are coupled by the conceptual design (Gausemeier *et al.*, 2011). Pahl *et al.* link product planning to the *clarification of the task* in product development to underline the necessity of a content-related junction and a work-related integration of the two (Pahl *et al.*, 2007). Ulrich and Eppinger describe product planning as the initial phase or phase zero of product development (Ulrich and Eppinger, 2016).

The success of product planning depends on the utilization of the available data which – esp. for CPS – can be enormous. The process of *accessing*, *aggregating* and *analyzing* large amounts of data from multiple sources is called data analytics (DA). It enables companies to *extract knowledge* from data to *understand historical* and *predict future events* (Tyagi, 2003). Data analytics is based on mathematics, computer science, and business analysis techniques (Porter and Heppelmann, 2015). It can be divided into four types with increasing value and complexity: *descriptive*, *diagnostic*, *predictive* and *prescriptive* analytics (Steenstrup *et al.*, 2014).

As shown, product planning and data analytics are two established and independent research areas. Together, they span the new research area of usage data-driven product planning which still needs to be thoroughly researched. Studies show that companies increasingly make decisions based on data analysis results, seeing data analytics as a crucial building block for creating value; yet, most companies do not utilize usage data within product creation processes (Erwin *et al.*, 2015). Also, research on the topic is scarce. From these considerations, four research questions are derived:

1. *Main concepts: What concepts constitute usage data-driven product planning?*
2. *Advantages: Why should companies pursue usage data-driven product planning?*
3. *Success factors: What factors contribute to the success of usage data-driven product planning?*
4. *Challenges: What makes the implementation of usage data-driven product planning difficult?*

## 3 RESEARCH DESIGN

To answer the research questions, we conducted a systematic literature review (SLR), following the suggestions of Webster and Watson (2002), Brocke *et al.* (2009) and Rowley and Slack (2004). First, we prepared a list of relevant journals on the basis of the Financial Times Research Rank by Ormans (2016) and an additional extensive web search for topic-related journals (Webster and Watson, 2002). We included journals which focus on disciplines like “Product Innovation”, “Innovation Management”, “Engineering Management”, “R&D Management”, “Systems Engineering”, “Data and Knowledge Engineering”, “Data Science”, “Big Data”, “Marketing Management”, “Operations Research” etc. as usage data-driven product planning is an interdisciplinary field. All in all, the final list included 44 journals. Second, we created a concept map to identify important concepts and their synonyms (Rowley and Slack, 2004; Brocke *et al.*, 2009). We used the concept map to iteratively create and test search strings as a combination of the concepts in the concept map. Through iterative

improvement, the following search string was formed: (“*product development*” OR “*product design*” OR “*product engineering*”) AND (“*usage data*” OR “*operational data*” OR “*lifecycle data*” OR “*big data*” OR “*data analytics*” OR “*closed-loop*” OR “*feedback information*” OR “*feedback data*”). Since the number of logical connectors (OR and AND) was limited, we decided to not include the term “product planning” as it is not widely used and our tests provided better results for the alternative terms “product development”, “product design” and “product engineering”. Third, we searched for articles within each journal (from 2005 to 2020) in the Google Scholar Database using the literature research tool “Publish or Perish” (Harzing, 2007). The search resulted in 1870 articles which we screened on a title and abstract basis. While hundreds of articles addressed topics like “big data” or “data analytics” in different contexts, the abstracts of only 36 papers suggested a focus on usage data-driven product planning. After reading these papers and conducting a backward and forward search (Webster and Watson, 2002), we identified 12 papers which contained answers to our research questions; the other papers were sorted out. Fourth, we examined papers from conference proceedings. From prior research, we knew of 14 papers within the context of usage data-driven product planning. We used these to conduct a backward and forward search, resulting in a list of 76 papers. From these, 12 papers contained answers to our research questions. Finally, following the suggestions of our reviewers, we added 4 more articles.

In total, we identified 28 highly relevant articles. For the detailed analysis, we read each paper twice, extracted all relevant information concerning our research questions and loosely clustered them into thematic chunks. Thereafter, we iteratively improved the clustering by critically investigating each cluster, splitting clusters up and building new clusters. As the result of this iterative process, we obtained the main concepts, advantages, success factors and challenges of usage data-driven product planning.

## 4 RESULTS

### 4.1 Main Concepts

Usage data-driven product planning consists of six main concepts (see Figure 1):

1. The product’s sensors capture real-time readings of the product in its operating environment (Fathi *et al.*, 2011; Porter and Heppelmann, 2015; Hou and Jiao, 2020; Igba *et al.*, 2015; Chowdhery *et al.*, 2020). The data include user-generated data (capturing user behavior), product operating data (capturing product behavior) and environmental data (capturing environment behavior) (Hou and Jiao, 2020, 2020). In addition, further data like manual reports can also be collected (Chowdhery *et al.*, 2020). The data collection approach can either be reactive and thus be driven by concrete events (e.g. machine failures) or it can be proactive by collecting data on a large scale and analyzing it exploratorily (Holler *et al.*, 2016b; Holler *et al.*, 2017).
2. Using cyber-infrastructure (van Horn *et al.*, 2012), the captured usage data are fed back into the product creation process (Porter and Heppelmann, 2014; Jun *et al.*, 2007), where they are valuable in all stages (Holler *et al.*, 2017; Hou and Jiao, 2020; Jun *et al.*, 2007), but offer the highest value in the early stages like product planning as these are characterized by lots of uncertainties and the determination of lifecycle costs (Holler *et al.*, 2016b; Holler *et al.*, 2017). Here, the data are used to objectively quantify product performance and usage profiles (van Horn *et al.*, 2012) to find usage-centric improvements for the product under consideration (Holler *et al.*, 2016b; Holmström Olsson and Bosch, 2013; Jun *et al.*, 2007; Hou and Jiao, 2020; van Horn *et al.*, 2012).
3. To identify improvements, statistical analysis, data mining and machine learning techniques must be applied (Hou and Jiao, 2020; Igba *et al.*, 2015). The data analysis can (a) build upon the available data in a bottom-up approach (less effort and faster implementation) or (b) start with a predefined objective in a top-down approach (more effort, clear future-focus) (Wilberg *et al.*, 2017b).
4. The results of the data analysis enable developers to make decisions based on facts instead of assumptions only (Hou and Jiao, 2020; Chowdhery *et al.*, 2020), thus improving the decision-making process (Jun *et al.*, 2007; Wu *et al.*, 2020; Fathi *et al.*, 2011).
5. Identified improvements can be implemented in existing and future products (Jun *et al.*, 2007; Wilberg *et al.*, 2017b; Abramovici *et al.*, 2017; Wu *et al.*, 2020; Xu *et al.*, 2016; van Horn *et al.*, 2012). For future products, especially the development of new product generations is suited for the analysis of feedback data (Holmström Olsson and Bosch, 2013; Igba *et al.*, 2015; Fathi *et al.*, 2011; Chowdhery *et al.*, 2020).

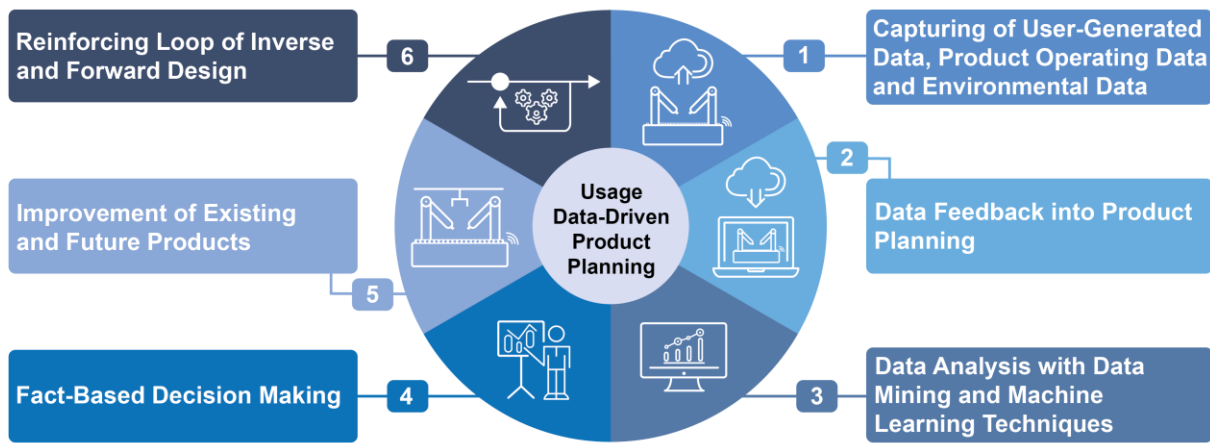


Figure 1: The six main concepts of usage data-driven product planning

6. By feeding back usage data into product planning, usage data-driven product planning represents an inverse design approach which will not replace, but complement the traditional, assumption-based forward design (Hou and Jiao, 2020; Wilberg *et al.*, 2017a). In conjunction, inverse and forward design form a reinforcing loop for the continuous and iterative improvement of product design (Hou and Jiao, 2020).

## 4.2 Advantages

Usage data-driven product planning offers a variety of advantages which are listed in a taxonomy in Table 1 and explained below. The advantages apply on three levels: **Analysis**, **Process** and **Business**.

Table 1. Taxonomy of the advantages of usage data-driven product planning

Level 1: Analysis	Level 2: Process	Level 3: Business
(1.1) Finding hidden information	(2.1) Improved customer- and user-involvement	(3.1) Improved decision-making processes
(1.2) Better product understanding	(2.2) Continuous requirements analysis	(3.2) Usage-centric product portfolio
(1.3) Better understanding of customer and user needs	(2.3) Reduction of hardware-prototyping and field-testing	(3.3) Higher delivery frequency of functionality
(1.4) Contextualize and evaluate qualitative and subjective data	(2.4) Faster product development	(3.4) Higher innovative strength

**Analysis-Level:** (1.1) The utilization of analytics approaches like data mining makes it possible to *find hidden information* within the data that would be impossible to find manually (Menon *et al.*, 2005). This is especially true when comparing many products and thousands of sensor-readings over time (Porter and Heppelmann, 2015). (1.2) The data lead to an *improved understanding of the product* in operation (Menon *et al.*, 2005; Holmström Olsson and Bosch, 2013). Usage data-driven product planning uses these insights to improve product design and thus eliminate failures instead of using data to predict them (Holler *et al.*, 2016a). (1.3) Analyzing usage data leads to a *better understanding of customer and user needs* and is likely to deliver better results than traditional approaches (Timoshenko and Hauser, 2019; Porter and Heppelmann, 2014). The data can be used to analyze customer behavior and preferences (Hou and Jiao, 2020; Porter and Heppelmann, 2015) and derive customer segments (Cantamessa *et al.*, 2020). For these, individual (Hou and Jiao, 2020; Holler *et al.*, 2016b), future (Holmström Olsson and Bosch, 2013), unspoken (Li *et al.*, 2015; Timoshenko and Hauser, 2019) and latent needs (Hou and Jiao, 2020; van Horn *et al.*, 2012) can be identified. (1.4) Furthermore, quantitative product usage data helps to *contextualize and evaluate qualitative and subjective data* (Holler *et al.*, 2016b).

**Process-Level:** (2.1) By building on usage data, the *customer and user involvement* in the development process is markedly *improved* (Hou and Jiao, 2020), leading to better collaboration

(Holler *et al.*, 2016b). (2.2) Feedback data enable a *continuous requirements analysis* (Timoshenko and Hauser, 2019) (e. g. through the continuous live elicitation of customer needs (Cantamessa *et al.*, 2020)) eliminating momentary biases. (2.3) As the feedback of usage data promotes a usage-centric and fact-based decision-making process, *hardware-prototyping and field-testing* necessities can be *reduced* significantly (Holler *et al.*, 2016b; Holler *et al.*, 2017). (2.4) Consequently, usage data-driven product planning allows for a *faster product development* process (Menon *et al.*, 2005).

**Business-Level:** (3.1) The integration of data analytics methods into product planning enables data- and fact-based decisions (Wuest *et al.*, 2014; Holler *et al.*, 2016a) while reducing assumption- and experience-based ones (Holler *et al.*, 2017; Holler *et al.*, 2016a), thus *improving decision-making processes* (Xu *et al.*, 2016; Menon *et al.*, 2005). (3.2) Analyzing usage data promotes a *usage-centric product portfolio* (Holler *et al.*, 2016a; Holler *et al.*, 2017; Holmström Olsson and Bosch, 2013). (3.3) The continuous requirement analysis and the faster development process help companies to *increase their delivery frequency of functionality* (Holmström Olsson and Bosch, 2013; Jun *et al.*, 2007) and thus react on insights not anticipated by the previous design (Cantamessa *et al.*, 2020). (3.4) Lastly, analyzing usage data enables companies to create *high-quality innovations* (Li *et al.*, 2015; Kiron *et al.*, 2014) as it pushes them towards new ideas at a higher speed (Erevelles *et al.*, 2016) and facilitates their acceptance by keeping the companies open-minded (Kiron *et al.*, 2014).

### 4.3 Success Factors

To exploit the advantages mentioned, the factors contributing significantly to the success of usage data-driven product planning must be identified. Table 2 shows a taxonomy of these success factors within four classes: **Organization**, **Product**, **Data Analysis** and **Evaluation**. The factors are described below.

Table 2. Taxonomy of the success factors for usage data-driven product planning

Class 1: Organization	Class 2: Product	Class 3: Data Analysis	Class 4: Evaluation
(1.1) Data strategy	(2.1) High number of intelligent components	(3.1) Standardized data and systems	(4.1) Comprehensible data analysis results
(1.2) Use cases	(2.2) Long operating times	(3.2) Effective data collection and feedback	(4.2) Decision support systems
(1.3) Ignorance-b. view & induct. reasoning	(2.3) High similarity with other products	(3.3) Proactive data collection and analysis	
(1.4) Cooperation of product & data experts	(2.4) Data access	(3.4) Joint analysis of heterogeneous data	
(1.5) Integration of DA with tradition. methods		(3.5) Complex data analyses	
(1.6) Turning consumers into prosumers		(3.6) Considering the product context	

**Organization:** (1.1) A *data strategy* is critical for success (Wilberg *et al.*, 2017a; Erwin *et al.*, 2015). It is the result of (a) an in-depth analysis of the product under consideration, the corporate strategy etc. and (b) a conceptual phase (Wilberg *et al.*, 2017b) which includes the definition of use cases among others (Wilberg *et al.*, 2017a). (1.2) *Use cases* create value by e. g. defining the analysis goal, asking open questions, outlining its potential benefits and deriving the data that need to be collected (Menon *et al.*, 2005; Wilberg *et al.*, 2017a; Wilberg *et al.*, 2017b). In order to identify opportunities and limitations, all relevant stakeholders must be involved (Wilberg *et al.*, 2017b). (1.3) An *ignorance-based view and inductive reasoning* promote the definition of a successful use case (Erevelles *et al.*, 2016). The ignorance-based view triggers questions with a high potential for hidden insights while inductive reasoning will create a greater understanding of the investigated object (Erevelles *et al.*, 2016). (1.4) As use cases require product knowledge, a close *cooperation of product and data experts* is crucial (Shahbaz *et al.*, 2006; Fathi *et al.*, 2011; Li *et al.*, 2015; Cantamessa *et al.*, 2020). Especially at the start of common projects, interaction points between product and data experts are numerous, making a direct communication and a common language a necessity (Li *et al.*, 2019). In the evaluation phase, multiple perspectives from a heterogeneous team may capture more

valuable insights from within the data (van Horn *et al.*, 2012). Hence, this so-called knowledge fusion is critical when evaluating use cases and analysis results (Xu *et al.*, 2016). (1.5) The *integration of data analytics with traditional methods* promises the highest new product success and knowledge fusion (Xu *et al.*, 2016; Hou and Jiao, 2020). (1.6) Lastly, for data feedback, companies need to renew their customer relationships by *turning consumers into prosumers* (Cantamessa *et al.*, 2020).

**Product:** (2.1) Products with a *high number of intelligent components* are especially suited for usage data-driven product planning as these generate the data necessary for a large scale data analysis (Holler *et al.*, 2016a; Jun *et al.*, 2007). (2.2) In order to generate enough data for the algorithms to work properly, *long operating times* are a prerequisite (Holler *et al.*, 2016a). (2.3) When comparing the data of multiple products, a *high similarity between these products* is essential (Holler *et al.*, 2016a). Due to this, usage data-driven product planning works best in product generation planning (Holler *et al.*, 2017) and when using the principles of design modularity and platforms (Cantamessa *et al.*, 2020). (2.4) Finally, *data access* must be achieved through providing customers transparency concerning data usage and incentives in form of a clear value proposition for sharing their data (Porter and Heppelmann, 2014; Wilberg *et al.*, 2017a).

**Data Analysis:** (3.1) *Standardized data and ecosystems* facilitate data exchange and analysis (Wilberg *et al.*, 2017a; Holler *et al.*, 2016a). For qualitative data, formal taxonomies and standardized information structures improve reuse (Goh and McMahan, 2009). (3.2) An *effective data collection and feedback process* is mandatory for usage data-driven product planning to work. It can be achieved through (a) smart sensing and advanced data-collection methods (Hou and Jiao, 2020) and through (b) clearly defined points of data capture and reuse as well as incentives within the people's work processes (Goh and McMahan, 2009; Menon *et al.*, 2005). The data must be mapped to concrete product instances as well as the product type (Abramovici *et al.*, 2017). (3.3) In general, a *proactive data collection and analysis* promises to deliver the most value (Holler *et al.*, 2016b). (3.4) In order to achieve valuable analysis results, *heterogeneous data* (e. g. structured and unstructured data) need to be *jointly analysed* (Porter and Heppelmann, 2015; Igba *et al.*, 2015; Abramovici *et al.*, 2017) in (3.5) *sophisticated data analyses* as these promise to achieve deeper insights and more satisfied users (Porter and Heppelmann, 2015; Erwin *et al.*, 2015; Cantamessa *et al.*, 2020). (3.6) Before analyzing data, it is important to *consider the context of the product and its data*: These contextual information need to be evaluated when deciding about the data analysis approach (Shahbaz *et al.*, 2006). Furthermore, they help to understand the analysis results, potentially leading to deeper insights (Wilberg *et al.*, 2017a).

**Evaluation:** (4.1) The *results* of the data analysis must be *easily comprehensible* as decision-makers are not involved in all analysis steps (Li *et al.*, 2015). Managers and product developers need to be presented aggregated and accurate knowledge instead of vast amounts of data (Abramovici *et al.*, 2017). (4.2) Furthermore, they need *decision support systems* which give them suitable advice (Jun *et al.*, 2007) and help them make deliberate decisions based on the information available.

#### 4.4 Challenges

While the success factors help companies to exploit the advantages of usage data-driven product planning, they will still face numerous challenges. Table 3 shows a taxonomy of challenges using the four classes **Organization**, **Product**, **Data Analysis** and **Evaluation**. The challenges are described below.

Table 3: Taxonomy of the challenges of usage data-driven product planning

Class 1: Organization	Class 2: Product	Class 3: Data Analysis	Class 4: Evaluation
(1.1) Definition of a data-strategy	(2.1) Highly individual products	(3.1) Selection of the data to be collected	(4.1) Validity-check of the analysis results
(1.2) Definition and selection of use cases	(2.2) Short-cyclical product improvements	(3.2) Overlooking data availability and usage	(4.2) Interpretation of data analysis results
(1.3) Positive cost-benefit-ratio	(2.3) Capturing of the product context	(3.3) Data pre-processing	(4.3) Creation of new ideas
(1.4) Integr. of DA into traditional processes	(2.4) Missing sensors	(3.4) Analysis of large, multimodal data	
	(2.5) Prohibition of data access	(3.5) Choice of analysis methods	

**Organization:** (1.1) *Defining a data-strategy* is a major challenge many companies struggle with (Erwin et al., 2015; Wilberg et al., 2017b). (1.2) This includes the *definition and selection of promising use cases* as initially it is often difficult to define use cases at all and later there may be too many to implement (Wilberg et al., 2017a; Wilberg et al., 2017b). (1.3) Use cases must offer a *positive cost-benefit-ratio* (Holler et al., 2016a; Wilberg et al., 2017b), remain within a company's capacities (Wilberg et al., 2017a) and show their usefulness (Holmström Olsson and Bosch, 2013; van Horn et al., 2012). (1.4) Another challenge is the *integration of data analytics into traditional product creation and decision-making processes* (Hou and Jiao, 2020; Wilberg et al., 2017b), both technically and organizationally (Wilberg et al., 2017a).

**Product:** (2.1) *Highly individual products* represent an obstacle as data analysis insights may not apply to other instances in the product class and thus may not be transferable (Holler et al., 2016a; Holler et al., 2017; Cantamessa et al., 2020). (2.2) Furthermore, *short-cyclical product improvements* contradict the necessity of long operating times within a given configuration (Holler et al., 2016a). (2.3) Facing these challenges and the reduced comparability of products, *capturing the product usage context* becomes even more important for data reuse (Igba et al., 2015), but it is a major challenge (Hou and Jiao, 2020; Menon et al., 2005). (2.4) Due to the added costs, products may also *miss sensors* and thus be unable to collect the data necessary (Wilberg et al., 2017a). When retrofitting sensors, functional and economic aspects must be considered (van Horn et al., 2012). (2.5) Moreover, customers may *prohibit data access* (Wilberg et al., 2017a), especially to sensitive data (van Horn et al., 2012), exposing companies to additional complexity and costs to obtain rights to the data (Porter and Heppelmann, 2014).

**Data Analysis:** (3.1) While sensors generate lots of data, user-generated data are also extensive for complex products (Timoshenko and Hauser, 2019). But as more collected data do not necessarily lead to better results (Hou and Jiao, 2020), the *selection of the data* to be collected must be consistent with the strategy (Porter and Heppelmann, 2014) and carefully planned (Menon et al., 2005; Hou and Jiao, 2020). (3.2) Here, another challenge arises: Companies often *lack an overview of data availability and usage* (Holmström Olsson and Bosch, 2013; Wilberg et al., 2017b) which prevents information reuse (Goh and McMahon, 2009). (3.3) Once the data are collected, *pre-processing* confronts companies with multiple obstacles. Data validity must be checked (Wilberg et al., 2017a) and multimodal data, structured and unstructured, must be transformed and integrated (Porter and Heppelmann, 2015; Hou and Jiao, 2020; Shahbaz et al., 2006; Abramovici et al., 2017) which can be overwhelming for companies (Kiron et al., 2014). Here, especially textual data are difficult to work with because of their heterogeneity (Menon et al., 2005). (3.4) In complex industries, *analyzing large, multimodal data* is necessary to investigate the product from diverse perspectives (Xu et al., 2016; Hou and Jiao, 2020). However, companies struggle with a substantial lack of experience (Wilberg et al., 2017b) and seldom use complex analyses (Erwin et al., 2015). (3.5) Lastly, the *choice and appropriateness of data analysis methods* largely affect the quality of the solutions (Cantamessa et al., 2020).

**Evaluation:** (4.1) Analyzed data forms the basis for product improvement. But before impactful decisions are made, the *validity of the data analysis results* must be checked (van Horn et al., 2012) as the data samples analyzed are typically limited (e. g. due to small sample sizes) (Hou and Jiao, 2020). Here, even small errors could lead to major mistakes in the solution (Hou and Jiao, 2020). Therefore, it is necessary that product experts verify data quality and analysis results (Li et al., 2019). (4.2) The *interpretation of the data analysis results* shall turn data into valuable insights. However, as inverse problems do not have a unique solution (Hou and Jiao, 2020), the correct interpretation represents a major challenge (Hou and Jiao, 2020; Xu et al., 2016). To extract valuable insights, knowledge in both engineering and data analytics are required. Also, contextual information needs to be considered to avoid wrong interpretations (Wilberg et al., 2017a). (4.3) Lastly, the *creation of new ideas* based on the analysis and interpretation of existing data is not trivial (Wu et al., 2020). Also, the derivation of improvements for a whole product class from (partially) subjective information about one instance is difficult (Abramovici et al., 2017).

## 5 DISCUSSION AND CONCLUSION

### 5.1 Key Insights

1. Usage data-driven product planning consists of **six main concepts** (see section 4.1). (1) The captured user-generated, product operating and environmental data (2) are fed back into product

planning and (3) analyzed with statistical analysis, data mining and machine learning techniques. (4) The results enable fact-based decisions (5) to improve existing and future products. (6) The feedback process represents an inverse design approach and forms a reinforcing loop with traditional forward design.

2. Usage data-driven product planning offers **advantages on three levels** (see section 4.2). On the analysis-level, the analysis of products, customers and users is improved. On the process-level, planning and development processes are enhanced while reducing time and costs. On the business-level, fact-based decisions enable a usage-centric product portfolio and a higher innovative strength.
3. **Success factors** and **challenges** can be structured using taxonomies with the classes organization, product, data analysis and evaluation (see sections 4.3 and 4.4). While some success factors and challenges are independent of each other, a significant number of success factors correspond directly with challenges: data strategy, use cases, integration of data analytics with traditional methods, long operating times, high similarity with other products, data access, effective data collection and feedback, joint analysis of heterogeneous data and complex data analyses. Their realization is decisive for the successful implementation of usage data-driven product planning.

## 5.2 Limitations

The limitations of our study are related to our research design. First, our journal selection process may have excluded further journals with relevant articles. Yet, we believe that by conducting a thorough backward and forward search, we found the most relevant articles. Second, due to high number of related concepts, the construction of our search string may not be optimal. We tested multiple options on a sample basis, but we did not investigate all papers for every option. Search strings including different concepts (like product planning, data mining) may yield results not found with our search string. Third, we searched only within the Google Scholar Database, setting up our dataset based on titles, abstracts, and keywords. Fourth, the selection of relevant articles as well as the clustering process were subjective. Other scholars might have included different articles and clustered the elements in another way.

## 5.3 Implications for Managers and Future Research

**Managers** can use the six main concepts to create a better understanding of the topic within their companies. Likewise, the listed advantages can be utilized to achieve a high level of acceptance and promote future projects. Using the success factors taxonomy, managers can assess the current and needed maturity of their companies for usage data-driven product planning (e. g. with a capability maturity model). Lastly, managers should analyze the challenges and create plans to overcome them.

As usage data-driven product planning is a new research area, various **future research** is needed. From our results, three research needs stand out: (1) The creation of a data strategy with promising use cases is crucial as it guides all following steps. Here, the data strategy as well as the use cases must align with the companies' corporate strategy and its current situation. (2) The integration of data analytics into traditional forward design processes is also critical as usage data-driven product planning as a form of inverse design will not replace but complement traditional forward design. (3) The utilization of data analytics must be made accessible to small and medium-sized enterprises. In contrast to large companies, they mostly lack resources to build up their own sophisticated solutions. Finally, as our research focuses on product usage data only, the integration of non-usage data should be analysed, e.g., *how can data from sources like customer reviews complement the usage phase data?* Especially for the identification of customer needs, there is already a high number of approaches using text mining techniques on social media and online reviews (Bertoni, 2020).

## ACKNOWLEDGEMENTS

The research and development project *DizRuPt* is funded by the German Federal Ministry of Education and Research (BMBF) within the "Innovations for Tomorrow's Production, Services and Work" Program. The author is responsible for the content of this publication. Thanks to the reviewers for their detailed feedback and their valuable literature suggestions.



## REFERENCES

- Abramovici, M., Gebus, P., Göbel, J.C. and Savarino, P. (2017), "Utilizing Unstructured Feedback Data from MRO Reports for the continuous improvement of standard products", *21st International Conference on Engineering Design (ICEDI17)*, Vancouver, Aug. 21-25, 2017, The Design Society, Glasgow, pp. 327–336.
- Bertoni, A. (2020), "Data-Driven Design in Concept Development: Systematic Review and Missed Opportunities", *Proceedings of the Design Society: DESIGN Conference*, Vol. 1, pp. 101–110. <https://doi.org/10.1017/dsd.2020.4>
- Brocke, J., Simons, A., Niehaves, B., Niehaves, B., Reimer, K., Plattfaut, R. and Cleven, A. (2009), "Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process", *European Conference on Information Systems (ECIS)*, Verona, Italy, June 8-10, 2009, pp. 2206–2217.
- Cantamessa, M., Montagna, F., Altavilla, S. and Casagrande-Seretti, A. (2020), "Data-driven design: the new challenges of digitalization on product design and development", *Design Science*, Vol. 6. <https://doi.org/10.1017/dsj.2020.25>.
- Chowdhery, S.A., Bertoni, M., Wall, J. and Larsson, T. (2020), "A data-driven design framework for early stage PSS design exploration", *Design Science*.
- Erevelles, S., Fukawa, N. and Swayne, L. (2016), "Big Data consumer analytics and the transformation of marketing", *Journal of Business Research*, Vol. 69, No. 2, pp. 897–904. <https://doi.org/10.1016/j.jbusres.2015.07.001>.
- Erwin, T., Heidkamp, P. and Pols, A. (2015), *Creating Value With Data: Report 2015*. KPMG and Bitkom Research. Available at: <https://www.bitkom.org/sites/default/files/file/import/KPMG-Bitkom-Research-Studie-MDWS-final-2.pdf> (March 3, 2020).
- Fathi, M., Abramovici, M., Holland, A., Lindner, A. and Dienst, S. (2011), "Usage scenarios of a knowledge-based assistance system for decision support in product improvement", *6th Conference on Professional Knowledge Management - From Knowledge to Action*, Innsbruck, Austria, February 21-23, 2011, Gesellschaft für Informatik e.V., Bonn, Germany, pp. 295–304.
- Gausemeier, J., Dumitrescu, R., Kahl, S. and Nordsiek, D. (2011), "Integrative development of product and production system for mechatronic products", *Robotics and Computer-Integrated Manufacturing*, Vol. 27 No. 4, pp. 772–778. <https://doi.org/10.1016/j.rcim.2011.02.005>
- Goh, Y.M. and McMahan, C. (2009), "Improving reuse of in-service information capture and feedback", *Journal of Manufacturing Technology Management*, Vol. 20 No. 5, pp. 626–639. <https://doi.org/10.1108/17410380910961028>.
- Harzing, A.W. (2007), *Publish or Perish*. Available at: <https://harzing.com/resources/publish-or-perish> (November 12, 2020).
- Holler, M., Neiditsch, G., Uebernickel, F. and Brenner, W. (2017), "Digital Product Innovation in Manufacturing Industries - Towards a Taxonomy for Feedback-driven Product Development Scenarios", *50th Hawaii International Conference on System Sciences (HICSS)*, Waikoloa Village, Hawaii, USA, January 4-7, 2017, pp. 4726-4735. <https://doi.org/10.21256/zhaw-3365>.
- Holler, M., Stoekli, E., Uebernickel, F. and Brenner, W. (2016a), "Towards Understanding closed-loop PLM: The Role of Product Usage Data for Product Development enabled by intelligent Properties", *29th Bled eConference on Digital Economy (Bled)*, Bled, Slovenia, June 19-22, 2016, Association for Information Systems, AIS Electronic Library, pp. 479–491.
- Holler, M., Uebernickel, F. and Brenner, W. (2016b), "Understanding the Business Value of Intelligent Products for Product Development in Manufacturing Industries", *8th International Conference on Information Management and Engineering (ICIME)*, Istanbul, Turkey, November 2-5, 2016, Association for Computing Machinery, New York, USA, pp. 18–24. <https://doi.org/10.1145/3012258.3012266>.
- Holmström Olsson, H. and Bosch, J. (2013), "Towards Data-Driven Product Development: A Multiple Case Study on Post-deployment Data Usage in Software-Intensive Embedded Systems", *4th International Conference on Lean Enterprise Software and Systems (LESS)*, Galway, Ireland, December 1-4, 2013, Springer, Berlin, Heidelberg, Germany, pp. 152–164. [https://dx.doi.org/10.1007/978-3-642-44930-7\\_10](https://dx.doi.org/10.1007/978-3-642-44930-7_10).
- Hou, L. and Jiao, R.J. (2020), "Data-informed inverse design by product usage information: a review, framework and outlook", *Journal of Intelligent Manufacturing*, Vol. 31 No. 3, pp. 529–552. <https://doi.org/10.1007/s10845-019-01463-2>.
- Igba, J., Alemzadeh, K., Gibbons, P.M. and Henningsen, K. (2015), "A framework for optimising product performance through feedback and reuse of in-service experience", *Robotics and Computer-Integrated Manufacturing*, Vol. 36, pp. 2–12. <https://doi.org/10.1016/j.rcim.2014.12.004>.
- Jun, H.-B., Shin, J.-H., Kiritsis, D. and Xirouchakis, P. (2007), "System architecture for closed-loop PLM", *International Journal of Computer Integrated Manufacturing*, Vol. 20 No. 7, pp. 684–698. <https://doi.org/10.1080/09511920701566624>.
- Kiron, D., Kirk Prentice, P. and Boucher Ferguson, R. (2014), "The Analytics Mandate", *MIT Sloan Management Review*, Vol. 55 No. 4, pp. 1–25.

- Lee, E.A. (2008), “Cyber Physical Systems: Design Challenges”, *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*, Orlando, Florida, USA, May 5-7, 2008, IEEE Computer Society, Los Alamitos, California, USA, pp. 363–369. <https://doi.org/10.1109/ISORC.2008.25>.
- Li, J., Tao, F., Cheng, Y. and Zhao, L. (2015), “Big Data in product lifecycle management”, *The International Journal of Advanced Manufacturing Technology*, Vol. 81, pp. 667–684. <https://doi.org/10.1007/s00170-015-7151-x>.
- Li, Y., Roy, U. and Saltz, J.S. (2019), “Towards an integrated process model for new product development with data-driven features (NPD3)”, *Research in Engineering Design*, Vol. 30 No. 2, pp. 271–289. <https://doi.org/10.1007/s00163-019-00308-6>.
- Menon, R., Tong, L.H. and Sathiyakeerthi, S. (2005), “Analyzing Textual Databases using Data Mining to Enable Fast Product Development Processes”, *Reliability Engineering & System Safety*, Vol. 88 No. 2, pp. 171–180. <https://doi.org/10.1016/j.res.2004.07.007>.
- Ormans, L. (2016), *50 Journals used in FT Research Rank*. Financial Times. Available at: <https://www.ft.com/content/3405a512-5cbb-11e1-8f1f-00144feabdc0> (November 12, 2020).
- Pahl, G., Beitz, W., Feldhusen, J. and Grote, K.-H. (2007), *Engineering design: A systematic approach*, 3rd ed., Springer, London, UK. <https://doi.org/10.1007/978-1-84628-319-2>.
- Porter, M.E. and Heppelmann, J.E. (2014), “How Smart, Connected Products Are Transforming Competition”, *Harvard Business Review*, Vol. 92 No. 11, pp. 64–88.
- Porter, M.E. and Heppelmann, J.E. (2015), “How Smart, Connected Products Are Transforming Companies”, *Harvard Business Review*, Vol. 93 No. 10, pp. 97–114.
- Rowley, J. and Slack, F. (2004), “Conducting a literature review”, *Management Research News*, Vol. 27 No. 6, pp. 31–39. <https://doi.org/10.1108/01409170410784185>.
- Shahbaz, M., Srinivas, M., Harding, J.A. and Turner, M. (2006), “Product Design and Manufacturing Process Improvement Using Association Rules”, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, Vol. 220 No. 2, pp. 243–254. <https://doi.org/10.1243/095440506X78183>.
- Steenstrup, K., Sallam, R., Eriksen, L. and Jacobson, S. (2014), *Industrial Analytics Revolutionizes Big Data in the Digital Business*. Gartner Research G00264728.
- Timoshenko, A. and Hauser, J.R. (2019), “Identifying Customer Needs from User-Generated Content”, *Marketing Science*, Vol. 38 No. 1, pp. 1–20. <https://doi.org/10.1287/mksc.2018.1123>.
- Tyagi, S. (2003), “Using data analytics for greater profits”, *Journal of Business Strategy*, Vol. 24 No. 3, pp. 12–14. <https://doi.org/10.1108/02756660310734938>.
- Ulrich, K.T. and Eppinger, S.D. (2016), *Product design and development*, 6th ed., McGraw-Hill, New York.
- van Horn, D., Olewnik, A. and Lewis, K. (2012), “Design Analytics: Capturing, Understanding, and Meeting Customer Needs Using Big Data”, *Proceedings of the ASME 2012 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, Chicago, Illinois, USA, August 12-15, 2012, American Society of Mechanical Engineers, pp. 863–875. <https://doi.org/10.1115/DETC2012-71038>.
- Webster, J. and Watson, R. (2002), “Analyzing the past to prepare for the future: Writing a literature review”, *Management Information Systems Quarterly*, Vol. 26 No. 2, pp. xiii–xxiii.
- Wilberg, J., Schäfer, F., Kandlbinder, P., Hollauer, C., Omer, M. and Lindemann, U. (2017a), “Data Analytics in Product Development: Implications from Expert Interviews”, *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, December 10-13, 2017, IEEE, pp. 818–822. <https://doi.org/10.1109/IEEM.2017.8290005>.
- Wilberg, J., Triep, I., Hollauer, C. and Omer, M. (2017b), “Big Data in product development: Need for a data strategy”, *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, Oregon, USA, July 9-13, 2017, IEEE, pp. 1–10. <https://doi.org/10.23919/PICMET.2017.8125460>.
- Wu, L., Hitt, L. and Lou, B. (2020), “Data Analytics, Innovation, and Firm Productivity”, *Management Science*, Vol. 66 No. 5, pp. 2017–2039. <https://doi.org/10.1287/mnsc.2018.3281>.
- Wuest, T., Hribernik, K. and Thoben, K.-D. (2014), “Capturing, Managing and Sharing Product Information along the Lifecycle for Design Improvement”, *10th International Workshop on Integrated Design Engineering*, Gommern, Germany, September 10-12, 2014, Chair of Information Technologies in Mechanical Engineering, Otto-von-Guericke-University, Magdeburg, Germany, pp. 107–115.
- Xu, Z., Frankwick, G.L. and Ramirez, E. (2016), “Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective”, *Journal of Business Research*, Vol. 69 No. 5, pp. 1562–1566. <https://doi.org/10.1016/j.jbusres.2015.10.017>.